

General Green's function formalism for layered systems: Wave function approach

Shu-Hui Zhang¹, Wen Yang^{1,*} and Kai Chang^{2,3†}

¹Beijing Computational Science Research Center, Beijing 100193, China

²SKLSM, Institute of Semiconductors, Chinese Academy of Sciences, P.O. Box 912, Beijing 100083, China and

³Synergetic Innovation Center of Quantum Information and Quantum Physics, University of Science and Technology of China, Hefei, Anhui 230026, China

The single-particle Green's function (GF) of mesoscopic structures plays a central role in mesoscopic quantum transport. The recursive GF technique is a standard tool to compute this quantity numerically, but it lacks physical transparency and is limited to relatively small systems. Here we present a numerically efficient and physically transparent GF formalism for a general layered structure. In contrast to the recursive GF that directly calculates the GF through the Dyson equations, our approach converts the calculation of the GF to the generation and subsequent propagation of a scattering wave function emanating from a local excitation. This viewpoint not only allows us to reproduce existing results in a concise and physically intuitive manner, but also provides analytical expressions of the GF in terms of a generalized scattering matrix. This identifies the contributions from each individual scattering channel to the GF and hence allows this information to be extracted quantitatively from dual-probe STM experiments. The simplicity and physical transparency of the formalism further allows us to treat the multiple reflection analytically and derive an analytical rule to construct the GF of a general layered system. This could significantly reduce the computational time and enable quantum transport calculations for large samples. We apply this formalism to perform both analytical analysis and numerical simulation for the two-dimensional conductance map of a realistic graphene p - n junction. The results demonstrate the possibility of observing the spatially-resolved interference pattern caused by negative refraction and further reveal a few interesting features, such as the distance-independent conductance and its quadratic dependence on the carrier concentration, as opposed to the linear dependence in uniform graphene.

PACS numbers: 73.23.Ad, 73.63.-b, 73.40.-c, 72.80.Vp

I. INTRODUCTION

The single-particle retarded Green's function (GF) is a key tool to calculate local and transport properties in mesoscopic systems^{1,2}, such as conductance, shot noise³, local density of states, and local currents^{4,5}. In the most popular scheme in which a scatterer is connected to two (or more) semi-infinite ballistic leads, the Landauer-Büttiker formula⁶⁻⁸ expresses the conductance $\sigma = (e^2/h)T(E_F)$ in terms of the transmission probability $T(E_F)$ across the scatterer on the Fermi surface. Typically, the electronic structure and transport properties of a mesoscopic system are described by a lattice model with a localized basis in real space, e.g., discretization of the continuous model⁹, empirical tight-binding¹⁰, or first-principles density-functional theory with a localized basis set^{11,12}. Then $T(E_F)$ is constructed from the lattice GF $G(E_F)$ across the scatterer through either an expression derived by Caroli *et al.*¹³ or the Fisher-Lee relations¹⁴⁻¹⁷ that express the scattering matrix $S(E_F)$ of the scatterer in terms of the lattice GF $G(E_F)$.

The recursive GF method (RGF) is a standard tool to compute the lattice GF of a scatterer connected to multiple leads^{2,18}. This method is reliable, computationally efficient, and allows for a parallel implementation¹⁹. It was pioneered by Thouless and Kirkpatrick²⁰ and by Lee and Fisher²¹. Then MacKinnon presented a "slice" formulation for a general layered system, which is the form most used nowadays²². Variations of the method were also introduced to treat multiple leads²³, arbitrary geometries^{24,25}, and local scatterers inside an infinite periodic system^{26,27}. With the development of many numerical algorithms, such as fast recursive or iterative schemes²⁸⁻³⁹ and closed-form solutions⁴⁰⁻⁴⁴,

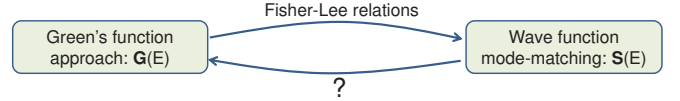


FIG. 1. Connection between two widely used approaches to mesoscopic quantum transport: the GF approach and wave function mode matching approach. The former calculates the GF $G(E)$, while the latter calculates the scattering matrix $S(E)$. The Fisher-Lee relations allow $S(E)$ to be constructed from $G(E)$, while the inverse relation remains absent for a general lattice model.

the development of RGF has culminated in many packages with different focus⁴⁵⁻⁴⁹ and is applicable to an arbitrary lattice Hamiltonian⁵⁰. However, these techniques and our knowledge about the lattice GF still suffer from two drawbacks/limitations.

First, there are two widely used approaches in mesoscopic quantum transport: the GF approach^{1,51} that computes the GF $G(E)$ and the wave function mode matching approach⁵²⁻⁵⁴ that computes the unitary scattering matrix $S(E)$. However, the connection between the GF $G(E)$ and $S(E)$ and hence the connection between these two approaches remain incomplete. It is well known that $S(E)$ can be constructed from $G(E)$ through the Fisher-Lee relations (see Fig. 1), as first derived by Fisher and Lee¹⁴ for a single parabolic band and two-terminal structures and later generalized to multiple leads¹⁵⁻¹⁷ and arbitrary lattice models^{25,41,55}. However, the inverse of this relation is nontrivial⁵⁶: explicit expressions of $G(E)$ in terms of $S(E)$ were derived¹⁵⁻¹⁷ only for a single parabolic band and in regions far from the scatterer. Generalization of this result to a general lattice model and over other regions would not only

complete the equivalence⁵⁵ between the GF approach and the wave function mode matching approach (Fig. 1), but also provides important tools to analyze the multi-probe scanning tunneling microscopy (STM), which has been applied to characterize a wide range of systems (see Refs. 57 and 58 for recent reviews) in the past few years, including nanowires^{59–61}, carbon nanotubes⁶², graphene nanoribbons^{63,64}, monolayer and bilayer graphene^{65–67}, and grain boundaries in graphene^{68,69} and copper⁷⁰. With one STM probe at \mathbf{R}_1 and the other STM probe at \mathbf{R}_2 , the Landauer-Büttiker formula expresses the conductance between the STM probes in terms of the GF $G(\mathbf{R}_2, \mathbf{R}_1, E_F)$, which provides spatially resolved information about the sample; e.g., with an analytical expression for the GF of pristine graphene^{26,27}, Settnes *et al.*⁷¹ were able to identify the different scattering processes of local scatterers in graphene. However, this analysis is still qualitative. To go one step further to extract *quantitatively* the information about the scatterers, an explicit expression of the GF in terms of the scattering matrix is required.

Second, the time cost of RGF increases rapidly with the number of localized basis required to subtend the sample. This imposes a computational limit when addressing realistic experimental samples; e.g., many quantum transport studies on graphene consider narrow graphene “nanoribbons” rather than large-area graphene. Three methods have been proposed to lift this constraint. The modular RGF^{17,36,72} is limited to electrons in a single parabolic band and specific shape of the sample. The other two methods essentially reduce the number of transverse bases, either by projecting the system Hamiltonian onto a small number of transverse modes^{73–75} or by assuming translational invariance⁷⁶ along the transverse direction. They could significantly reduce the time cost for wide samples, but the time cost still increases linearly with the length of the scatterer (along which transport occurs). To study a *long* sample, a more efficient method is desirable.

The origin of the above drawbacks/limitations is probably that the RGF treats the GF as a matrix and constructs the GF by a series of matrix recursion rules derived from the Dyson equation. Interestingly, although the rules for constructing the scattering matrix in terms of the GF (i.e., the Fisher-Lee relation) are concise and physically intuitive, their rigorous derivation (in which the GF is treated as a matrix) turns out to be rather tedious (see, e.g., Refs. 15, 16, and 25). This somewhat surprising fact suggests the possible existence of a very different way to represent and calculate the GF. This could not only enable a straightforward physical interpretation of the final results, but also shed light on some previous debates^{53,55,77} on the relationship between different calculation techniques in mesoscopic quantum transport.

In this work, we develop a numerically efficient and physically transparent GF formalism to address the above issues in layered systems, i.e., any system that is non-periodic along one direction, but is finite or periodic along the other directions. This includes a wide range of physical systems, such as interfaces and junctions, Hall bars, nanowires, multilayers, superlattices, carbon nanotubes, and graphene nanoribbons. Compared with the RGF that directly calculates the GF as a matrix through the Dyson equations, our approach

converts the calculation of the GF to the generation and subsequent propagation of a scattering wave function emanating from a local excitation. This viewpoint provides several advantages. First, the procedures for calculating the GF $G(E)$ becomes physically transparent and existing results from the RGF (such as the Fisher-Lee relation) can be derived in a concise and physically intuitive manner. Second, the GF $G(E)$ can be readily expressed in terms of a few scattering wave functions with energy E . This provides an *on-shell* generalization of the standard spectral expansion in classic textbooks on quantum mechanics^{78–80}, $G(\mathbf{R}_2, \mathbf{R}_1, E) = \sum_\lambda \langle \mathbf{R}_1 | \Psi_\lambda \rangle \langle \Psi_\lambda | \mathbf{R}_2 \rangle / (E + i0^+ - E_\lambda)$, which involves *all* the eigenstates $\{|\Psi_\lambda\rangle\}$ and eigenenergies $\{E_\lambda\}$ of the system. In terms of a generalized scattering matrix $S(E)$ that describes the scattering of both traveling and evanescent waves, we further establish a one-to-one correspondence between $G(E)$ and $S(E)$ (see Fig. 1). This identifies the contributions from each individual scattering channel (including evanescent channels) to the GF and hence allows this information to be extracted quantitatively from dual-probe STM. Third, the simplicity and physical transparency of the formalism further allows us to perform an infinite summation of the multiple reflection between different scatterers and arrive at an analytical construction rule for the GF of a general layered system containing an arbitrary number of scatterers. This could make the time cost independent of the length of the sample along the transport direction and hence significantly speed up the calculation. By further reducing the number of bases along the transverse direction^{73–76}, our formalism enables quantum transport calculations over macroscopic distances and on large samples.

Recently, the chiral tunneling^{81–83} and negative refraction^{84–86} of graphene p - n junctions have received a lot of interest and the anomalous focusing effect was observed experimentally^{87,88}, but previous theoretical studies are mostly based on the low-energy continuous model, whose validity is limited to the vicinity of the Dirac points. Here we apply our GF approach to perform an analytical analysis and numerical simulation for the two-dimensional conductance map of dual-probe STM experiments in a realistic graphene p - n junction described by the tight-binding model. The results demonstrate the possibility of observing the spatially resolved interference pattern caused by negative refraction and further reveals some interesting features (such as the distance-independent conductance and its quadratic dependence on the carrier concentration, as opposed to the linear dependence in uniform graphene) that may also be observed in dual-probe STM experiments.

This paper is organized as follows. In Sec. II, we introduce the model, review the commonly used RGF technique, and presents the key idea of our approach. In Sec. III, we derive the GF of an infinite system containing a single scatterer, as well as an analytical construction rule for the GF of a general layered system containing an arbitrary number of scatterers. In Sec. IV, we express the GF analytically in terms of a generalized scattering matrix or in terms of a few scattering states on the energy shell E . In Sec. V, we exemplify our results in a 1D chain and then apply it to analyze and simulate the

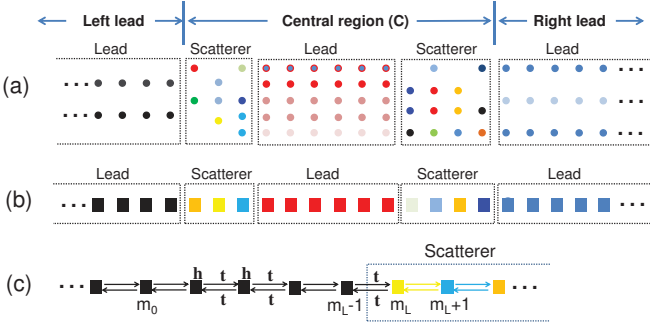


FIG. 2. (a) A layered 2D structure consisting of multiple periodic slices (i.e., leads) and disordered slices (i.e., scatterers). (b) Regarding each slice as a unit cell (filled squares), the structure in (a) becomes a 1D lattice. (c) A semi-infinite lead connected to a scatterer. The unit cell Hamiltonian (filled squares) and nearest-neighbor hopping (double arrows) are m -independent inside the lead, but could be disordered inside the scatterer.

two-dimensional conductance map of a realistic graphene p - n junction. Finally, a brief conclusion is given in Sec. VI.

II. THEORETICAL MODEL AND KEY IDEAS

We consider a general layered system in the lattice representation. When each layer is an infinite, periodic repetition of a basic unit, we can make a Fourier transform to effectively reduce each layer to a single basic unit. Disorder inside each layer can also be introduced by using a sufficiently large unit cell and repeating it periodically. Then we can regard each layer as a finite-size unit cell, so the system becomes a 1D lattice, e.g., by taking each layer/slice of the structure in Fig. 2(a) as a unit cell, Fig. 2(a) becomes Fig. 2(b). A general 1D lattice can always be decomposed into a few nonperiodic regions (referred to as *scatterers*) consisting of different unit cells sandwiched between periodic regions (referred to as *leads*) consisting of identical unit cells; see Fig. 2(b) for an example.

Without losing generality, we consider nearest-neighbor hopping⁸⁹ and use M_m to denote the number of orthonormal local bases in the m th unit cell. In the representation of these bases, the lattice Hamiltonian is an infinite-dimensional block-tridiagonal matrix:

$$\mathbf{H} = \begin{bmatrix} \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \mathbf{H}_{-2,-2} & \mathbf{H}_{-2,-1} & 0 & 0 & 0 & \cdots \\ \cdots & \mathbf{H}_{-2,-1}^\dagger & \mathbf{H}_{-1,-1} & \mathbf{H}_{-1,0} & 0 & 0 & \cdots \\ \cdots & 0 & \mathbf{H}_{-1,0}^\dagger & \mathbf{H}_{0,0} & \mathbf{H}_{0,1} & 0 & \cdots \\ \cdots & 0 & 0 & \mathbf{H}_{0,1}^\dagger & \mathbf{H}_{1,1} & \mathbf{H}_{1,2} & \cdots \\ \cdots & 0 & 0 & 0 & \mathbf{H}_{1,2}^\dagger & \mathbf{H}_{2,2} & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix}, \quad (1)$$

consisting of the $M_m \times M_{m+1}$ hopping matrix $\mathbf{H}_{m,m+1}$ between neighboring unit cells, its Hermitian conjugate $\mathbf{H}_{m+1,m} = \mathbf{H}_{m,m+1}^\dagger$, and the $M_m \times M_m$ Hamiltonian matrix $\mathbf{H}_{m,m}$ of the m th unit cell. In a lead, $\mathbf{H}_{m,m} = \mathbf{h}$ and $\mathbf{H}_{m,m+1} = \mathbf{t}$ are independent of m . In a scatterer, $\mathbf{H}_{m,m}$ and $\mathbf{H}_{m,m+1}$ could depend

on m arbitrarily. Here, as a convention, the region of the scatterers is chosen such that the hopping between the lead and the surface of a scatterer is *the same* as that inside this lead, e.g., $\mathbf{H}_{m_L-1,m_L} = \mathbf{t}$ in Fig. 2(c), where \mathbf{t} is the hopping inside the left lead. Except for the Hermiticity requirement $\mathbf{H} = \mathbf{H}^\dagger$, the lattice Hamiltonian is completely general. The (retarded) GF of the layered system is an infinite-dimensional matrix: $\mathbf{G}(E) \equiv (z - \mathbf{H})^{-1}$, where $z \equiv E + i0^+$. The GF from the unit cell m_0 to the unit cell m is an $M_m \times M_{m_0}$ matrix and corresponds to the (m, m_0) block of $\mathbf{G}(E)$, i.e., $\mathbf{G}_{m,m_0}(E) \equiv [(z - \mathbf{H})^{-1}]_{m,m_0}$. Hereafter we consider a fixed energy E or $z \equiv E + i0^+$ and omit this argument for brevity.

To highlight the distinguishing features of our approach and introduce relevant concepts, we first review the commonly used RGF method before presenting our idea.

A. Recursive Green's function method

The idea of the RGF is to build up the entire system out of disconnected subsystems by the Dyson equation. Let us start from two disconnected subsystems A and B characterized by the Hamiltonian $\mathbf{H}^{(A)}$ and $\mathbf{H}^{(B)}$, respectively. The (retarded) GFs of each subsystem are $\mathbf{G}^{(A)} \equiv (z - \mathbf{H}^{(A)})^{-1}$ and $\mathbf{G}^{(B)} \equiv (z - \mathbf{H}^{(B)})^{-1}$. Next we connect the interface (denoted by a) of A and the interface (denoted by b) of B by local couplings \mathbf{V}_{ab} and \mathbf{V}_{ba} . Thus the Dyson equation gives the GF

$$\mathbf{G} = \begin{bmatrix} \mathbf{G}_{AA} & \mathbf{G}_{AB} \\ \mathbf{G}_{BA} & \mathbf{G}_{BB} \end{bmatrix} \quad (2)$$

of the connected system in terms of the GFs of each subsystem⁹⁰:

$$\mathbf{G}_{AA} = ((\mathbf{G}^{(A)})^{-1} - \mathbf{V}_{ab}\mathbf{G}_{bb}^{(B)}\mathbf{V}_{ba})^{-1}, \quad (3a)$$

$$\mathbf{G}_{BB} = ((\mathbf{G}^{(B)})^{-1} - \mathbf{V}_{ba}\mathbf{G}_{aa}^{(A)}\mathbf{V}_{ab})^{-1}, \quad (3b)$$

$$\mathbf{G}_{BA} = \mathbf{G}_{Bb}^{(B)}\mathbf{V}_{ba}\mathbf{G}_{aA}, \quad (3c)$$

$$\mathbf{G}_{AB} = \mathbf{G}_{Aa}^{(A)}\mathbf{V}_{ab}\mathbf{G}_{bB}, \quad (3d)$$

or vice versa:

$$\mathbf{G}^{(A)} = \mathbf{G}_{AA} - \mathbf{G}_{Ab}\mathbf{V}_{ba}(1 + \mathbf{G}_{ab}\mathbf{V}_{ba})^{-1}\mathbf{G}_{aA}, \quad (4a)$$

$$\mathbf{G}^{(B)} = \mathbf{G}_{BB} - \mathbf{G}_{Ba}\mathbf{V}_{ab}(1 + \mathbf{G}_{ba}\mathbf{V}_{ab})^{-1}\mathbf{G}_{bB}. \quad (4b)$$

Equation (3) is the key to building up the entire system out of disconnected subsystems, while Eq. (4) can be used to calculate the GF of each subsystem when the GF of the connected system is known (e.g., if the connected system is infinite and periodic²⁷). The first two equations of Eq. (3) show that if we focus on one subsystem (say A), the presence of the other subsystem B amounts to a self-energy correction to the interface of A : $\mathbf{H}_{a,a}^{(A)} \rightarrow \mathbf{H}_{a,a}^{(A)} + \mathbf{V}_{ab}\mathbf{G}_{bb}^{(B)}\mathbf{V}_{ba}$.

1. General procedures of RGF

In RGF, to calculate the conductance of the general layered system as described earlier [Eq. (1)], the system is first partitioned into the semi-infinite left lead L (unit cells $m \leq 0$), the

central region C (unit cells $1 \leq m \leq N$), and the semi-infinite right lead R (unit cells $m \geq N + 1$). The *entire* central region C is regarded as a scatterer [see Fig. 2(a) for an example], so the GF is

$$\mathbf{G} = \begin{bmatrix} \mathbf{G}_{LL} & \mathbf{G}_{LC} & \mathbf{G}_{LR} \\ \mathbf{G}_{CL} & \mathbf{G}_{CC} & \mathbf{G}_{CR} \\ \mathbf{G}_{RL} & \mathbf{G}_{RC} & \mathbf{G}_{RR} \end{bmatrix}. \quad (5)$$

Let us use $\mathbf{H}^{(C)}$ for the Hamiltonian of the central region, and $\mathbf{H}^{(p)}$ ($p = L, R$) for the Hamiltonian of the lead p , as characterized by the unit cell Hamiltonian \mathbf{h}_p and nearest-neighbor hopping $\mathbf{t}_p = \mathbf{H}_{m,m+1}^{(p)}$. Then the central region part of the GF is computed from

$$\mathbf{G}_{CC} = (z - \mathbb{H})^{-1}, \quad (6)$$

where \mathbb{H} is the effective central region Hamiltonian: it equals $\mathbf{H}^{(C)}$ in the interior of C , but differs from $\mathbf{H}^{(C)}$ at the two surface unit cells:

$$\mathbb{H}_{1,1} = \mathbf{H}_{1,1} + \boldsymbol{\Sigma}^{(L)}, \quad (7a)$$

$$\mathbb{H}_{N,N} = \mathbf{H}_{N,N} + \boldsymbol{\Sigma}^{(R)}, \quad (7b)$$

due to self-energy corrections from the left and right leads:

$$\boldsymbol{\Sigma}^{(L)} = \mathbf{t}_L^\dagger \mathbf{G}_s^{(L)} \mathbf{t}_L, \quad (8)$$

$$\boldsymbol{\Sigma}^{(R)} = \mathbf{t}_R \mathbf{G}_s^{(R)} \mathbf{t}_R^\dagger. \quad (9)$$

Here $\mathbf{G}_s^{(L)} = [(z - \mathbf{H}^{(L)})^{-1}]_{0,0}$ is the GF of the left lead at the right surface, and $\mathbf{G}_s^{(R)} = [(z - \mathbf{H}^{(R)})^{-1}]_{N+1,N+1}$ is the GF of the right lead at the left surface, so they are referred to as *surface GFs* in the literature. Finally, to compute the linear conductance, we need to set $E = E_F$ and use the Landauer-Buttiker formula^{6–8} $\sigma = (e^2/h)T(E_F)$, where¹³

$$T(E_F) = \text{Tr} \mathbf{\Gamma}^{(R)} \mathbf{G}_{N,1} \mathbf{\Gamma}^{(L)} (\mathbf{G}_{N,1})^\dagger \quad (10)$$

and $\mathbf{\Gamma}^{(\alpha)} \equiv i(\boldsymbol{\Sigma}^{(\alpha)} - h.c.)$. Note that Eq. (10) only involves $\mathbf{G}_{N,1}$, the $(N, 1)$ block of \mathbf{G}_{CC} . Instead of direct matrix inversion [see Eq. (6)], $\mathbf{G}_{N,1}$ can be computed by building up the central region layer by layer² through Eq. (3). Let us use $\mathbf{G}^{(n)}$ ($n = 1, 2, \dots, N$) to denote the GF of the subsystem consisting of the unit cells $m = 1, 2, \dots, n$. The RGF starts from $\mathbf{G}^{(1)} = (z - \mathbb{H}_{1,1})^{-1}$, first uses the iteration

$$\mathbf{G}_{n,n}^{(n)} = (z - \mathbb{H}_{n,n} - \mathbf{H}_{n,n-1} \mathbf{G}_{n-1,n-1}^{(n-1)} \mathbf{H}_{n-1,n})^{-1} \quad (11)$$

to obtain $\{\mathbf{G}_{n,n}^{(n)}\}$, and then uses the iteration

$$\mathbf{G}_{1,n}^{(n)} = \mathbf{G}_{1,n-1}^{(n-1)} \mathbf{H}_{n-1,n} \mathbf{G}_{n,n}^{(n)} \quad (12)$$

to obtain $\mathbf{G}_{1,N} = \mathbf{G}_{1,N}^{(N)}$. The number of iterations and hence the time cost of the above recursive algorithm scales linearly with the length of the scattering region.

Equation (10) gives the total transmission probability, i.e., the sum of the transmission probabilities of all channels. To identify the contributions from each individual transmission

channels, it is necessary to use the GF to construct the transmission amplitude $S_{\beta,\alpha}^{(RL)}$ from the α th traveling channel in the lead L to the β th traveling channel in the lead R through the Fisher-Lee relations^{14–17,25,41,55} and then sum over all the traveling channels:

$$T(E_F) = \sum_{\alpha\beta \in \text{traveling}} |S_{\beta,\alpha}^{(RL)}|^2. \quad (13)$$

Alternatively, it is also possible to calculate the transmission amplitudes $\{S_{\beta,\alpha}^{(RL)}\}$ (and more generally the entire scattering matrix) by directly calculating the scattering of an incident traveling wave through the wave function mode matching approach^{52–54}. The equivalence between Eqs. (10) and (13), which establishes a connection between the GF approach and the wave function mode matching approach, is well known for a single parabolic band^{1,2,14–17}. For a general lattice model, there was suspicion⁷⁷ that Eq. (13) was incomplete since the GF in Eq. (10) includes both traveling waves and evanescent waves, while Eq. (13) only includes the contributions from traveling waves. Later, a rigorous equivalence proof was provided by Khomyakov *et al.*⁵⁵ and others²⁵, but the presence of the evanescent states does suggest that the GF is not completely equivalent to the unitary scattering matrix.

There are still two remaining issues: the calculation of the self-energies $\boldsymbol{\Sigma}^{(L,R)}$ (or equivalently the surface GFs $\mathbf{G}_s^{(L,R)}$) and a proper definition of the scattering channels and the transmission amplitudes $S_{\beta,\alpha}^{(RL)}$.

2. Self-energies: recursive method and eigenmode method

The numerical algorithms for computing $\boldsymbol{\Sigma}^{(L,R)}$ or equivalently the surface GFs $\mathbf{G}_s^{(L,R)}$ can be classified into two groups: recursive methods and eigenmode methods (see Ref. 90 for a review). The former calculates an approximate surface GF through some recursive relations, while the latter provides exact—within the numerical precision—closed-form solutions to the surface GF.

The idea of the recursive methods is to split the left lead into the surface unit cell $m = 0$ (subsystem A) and the remaining part (subsystem B); the Dyson equation [Eq. (3a)] gives the recursive relation

$$\begin{aligned} \mathbf{G}_s^{(L)} &= (z - \mathbf{h}_L - \mathbf{t}_L^\dagger \mathbf{G}_s^{(L)} \mathbf{t}_L)^{-1} \\ \Leftrightarrow \boldsymbol{\Sigma}^{(L)} &= \mathbf{t}_L^\dagger (z - \mathbf{h}_L - \boldsymbol{\Sigma}^{(L)})^{-1} \mathbf{t}_L. \end{aligned} \quad (14)$$

Similarly, by splitting the right lead into the surface unit cell $m = N + 1$ (subsystem A) and the remaining part (subsystem B), Eq. (3a) gives the recursive relation

$$\begin{aligned} \mathbf{G}_s^{(R)} &= (z - \mathbf{h}_R - \mathbf{t}_R \mathbf{G}_s^{(R)} \mathbf{t}_R^\dagger)^{-1} \\ \Leftrightarrow \boldsymbol{\Sigma}^{(R)} &= \mathbf{t}_R (z - \mathbf{h}_R - \boldsymbol{\Sigma}^{(R)})^{-1} \mathbf{t}_R^\dagger. \end{aligned} \quad (15)$$

Thus the surface GFs and self-energies can be obtained by simple or more efficient iteration techniques^{32,33}.

The eigenmode method has been derived independently several times^{40,41,46,52,53,77,91} and has been shown to be superior in accuracy and performance⁴⁰ compared to the recursive

methods. The central results are explicit expressions for the self-energies:

$$\Sigma^{(L)} = \mathbf{t}_L^\dagger (\mathbf{P}_-^{(L)})^{-1}, \quad (16a)$$

$$\Sigma^{(R)} = \mathbf{t}_R \mathbf{P}_+^{(R)}, \quad (16b)$$

and surface GFs:

$$\mathbf{G}_s^{(L)} = (z - \mathbf{h}_L - \mathbf{t}_L^\dagger (\mathbf{P}_-^{(L)})^{-1})^{-1} = (\mathbf{t}_L \mathbf{P}_-^{(L)})^{-1}, \quad (17)$$

$$\mathbf{G}_s^{(R)} = (z - \mathbf{h}_R - \mathbf{t}_R \mathbf{P}_+^{(R)})^{-1} = \mathbf{P}_+^{(R)} (\mathbf{t}_R^\dagger)^{-1}, \quad (18)$$

in terms of the (retarded) *propagation matrices* $\mathbf{P}_\pm^{(L,R)}$ (also referred to as Bloch matrices⁵⁵ or amplitude transfer matrices⁹⁰ in the literature), which can be constructed from the (retarded) *eigenmodes* of each lead.

Now we introduce the propagation matrices and eigenmodes in some detail, since they will play a central role in our GF approach. Let us consider a lead characterized by the unit cell Hamiltonian \mathbf{h} and nearest-neighbor hopping matrix $\mathbf{t} = \mathbf{H}_{m,m+1}$. The wave propagation in this lead is governed by the uniform Schrödinger equation

$$-\mathbf{t}^\dagger |\Phi(m-1)\rangle + (z - \mathbf{h}) |\Phi(m)\rangle - \mathbf{t} |\Phi(m+1)\rangle = 0, \quad (19)$$

where $z \equiv E + i0^+$. Imposing the Bloch symmetry $|\Phi(m)\rangle = e^{ikma} |\Phi\rangle$ (a is the thickness of each unit cell) gives

$$z |\Phi\rangle = (e^{-ika} \mathbf{t}^\dagger + \mathbf{h} + e^{ika} \mathbf{t}) |\Phi\rangle \equiv \mathbf{H}(k) |\Phi\rangle \quad (20)$$

for the eigenvector $|\Phi\rangle$. For an infinite lead, the wave function $|\Phi(m)\rangle$ must remain finite at $m \rightarrow \pm\infty$. This natural boundary condition dictates k to be real, so that Eq. (20) gives M real energy bands of the lead, where M is the number of basis states in each unit cell of this lead. For certain complex k 's, the energies could still be real, which form the complex energy bands of the lead.

Conversely, given the energy E and without imposing any boundary conditions, Eq. (19) or (20) can be solved to yield $2M$ (retarded) eigenmodes $\{k, |\Phi\rangle\}$ (see Appendix A)^{52–55}, where the wave vector k could be either real (i.e., *traveling* modes) or complex (i.e., *evanescent* modes). The eigenmodes are just the collection of eigenstates on the energy shell E in the real and complex energy bands of the lead. As a convention, each eigenvector $|\Phi\rangle$ should be normalized to unity, but different eigenvectors are not necessarily orthogonal. For a traveling eigenmodes with wave vector k and eigenvector $|\Phi\rangle$, its group velocity is

$$v = \partial_k \langle \Phi | \mathbf{H}(k) | \Phi \rangle = -2a \text{Im} \langle \Phi | \mathbf{t} e^{ika} | \Phi \rangle, \quad (21)$$

where $\langle \Phi |$ is the conjugate transpose of $|\Phi\rangle$, i.e., an M -component row vector. Then the $2M$ eigenmodes are classified into M right-going ones and M left-going ones: the former consist of traveling modes with a positive group velocity and evanescent modes decaying exponentially to the right (i.e., $\text{Im } k > 0$), while the latter consist of traveling modes with a negative group velocity and evanescent modes decaying exponentially to the left (i.e., $\text{Im } k < 0$). For clarity, we denote the M right-going eigenmodes as $\{k_{+, \alpha}, |\Phi_{+, \alpha}\rangle\}$ and the M left-going eigenmodes as $\{k_{-, \alpha}, |\Phi_{-, \alpha}\rangle\}$, where $\alpha = 1, 2, \dots, M$.

For every right-going evanescent mode $(+, \alpha)$ with wave vector $k_{+, \alpha}$, there is always a left-going evanescent mode $(-, \alpha)$ with wave vector $k_{-, \alpha} = k_{+, \alpha}^*$ ^{41,55,92}.

The propagation matrix \mathbf{P}_s for left-going ($s = -$) or right-going ($s = +$) waves is constructed as^{52–55}

$$\mathbf{P}_s \equiv \mathbf{U}_s \begin{bmatrix} e^{ik_{s,1}a} & & \\ & \ddots & \\ & & e^{ik_{s,M}a} \end{bmatrix} \mathbf{U}_s^{-1}, \quad (22)$$

where $\mathbf{U}_s \equiv [|\Phi_{s,1}\rangle, \dots, |\Phi_{s,M}\rangle]$ (i.e., its α th column is $|\Phi_{s,\alpha}\rangle$). The propagation matrices are standard tools in the wave function mode matching approach^{52–55} to describe wave propagation; e.g., a general right-going wave that obeys Eq. (19) can be written as $|\Phi(m)\rangle = \mathbf{P}_+^m |\Phi(0)\rangle$, while a general left-going wave obeying Eq. (19) can be written as $|\Phi(m)\rangle = \mathbf{P}_-^m |\Phi(0)\rangle$.

3. Scattering channels and Fisher-Lee relations

Since the GF $\mathbf{G}(E)$ describes the scattering of both traveling and evanescent eigenmodes by the central region, while the unitary scattering matrix $\mathbf{S}(E)$ describes the scattering of traveling eigenmodes only, it is possible to construct $\mathbf{S}(E)$ in terms of $\mathbf{G}(E)$, i.e., the Fisher-Lee-type relations. Compared with the Caroli's expression [Eq. (10)] that gives the total transmission probability, the Fisher-Lee relations further provide information about the scattering of each individual traveling eigenmode. For a general lattice model, different eigenmodes $\{|\Phi_{s,\alpha}\rangle\}$ are not orthogonal; then for each lead, it is necessary to introduce $2M$ (retarded) dual vectors $\{|\phi_{s,\alpha}\rangle\}$ through^{25,41,55}

$$\begin{bmatrix} \langle \phi_{s,1} | \\ \vdots \\ \langle \phi_{s,M} | \end{bmatrix} \equiv \mathbf{U}_s^{-1}, \quad (23)$$

where M is the number of bases in each unit cell of this lead. In general, different left-going (right-going) eigenvectors are not orthogonal, so \mathbf{U}_s is not necessarily unitary and $|\phi_{s,\alpha}\rangle$ is not necessarily equal to $|\Phi_{s,\alpha}\rangle$, but we always have the orthonormalization and completeness relations

$$\langle \phi_{s,\alpha} | \Phi_{s,\beta} \rangle = \langle \Phi_{s,\alpha} | \phi_{s,\beta} \rangle = \delta_{\alpha\beta}, \quad (24)$$

$$\sum_\alpha |\Phi_{s,\alpha}\rangle \langle \phi_{s,\alpha}| = \sum_\alpha |\phi_{s,\alpha}\rangle \langle \Phi_{s,\alpha}| = \mathbf{I}, \quad (25)$$

which follow from $\mathbf{U}_s \mathbf{U}_s^{-1} = \mathbf{U}_s^{-1} \mathbf{U}_s = \mathbf{I}$ (\mathbf{I} is the identity matrix). By inserting the completeness relation, any column vector $|\Phi\rangle$ can be expanded as a linear combination of either the M left-going eigenvectors or the M right-going eigenvectors as $|\Phi\rangle = \sum_\alpha c_{s,\alpha} |\Phi_{s,\alpha}\rangle$ with $c_{s,\alpha} = \langle \phi_{s,\alpha} | \Phi \rangle$. In terms of the eigenmodes and their dual vectors, Eq. (22) can be written as

$$\mathbf{P}_s \equiv \sum_\alpha e^{ik_{s,\alpha}a} |\Phi_{s,\alpha}\rangle \langle \phi_{s,\alpha}|, \quad (26)$$

which has a clear physical interpretation. For example, a general right-going wave is

$$|\Phi(m)\rangle = \mathbf{P}_+^m |\Phi(0)\rangle = \sum_\alpha e^{ik_{+, \alpha}ma} |\Phi_{+, \alpha}\rangle \langle \phi_{+, \alpha} | \Phi(0)\rangle, \quad (27)$$

i.e., $|\Phi(0)\rangle$ is first expanded as a linear combination of right-going eigenmodes $|\Phi(0)\rangle = \sum_{\alpha} |\Phi_{+, \alpha}\rangle \langle \phi_{+, \alpha} | \Phi(0)\rangle$ and then each right-going eigenmode propagates as $|\Phi_{+, \alpha}\rangle \rightarrow e^{ik_{+, \alpha} m a} |\Phi_{+, \alpha}\rangle$.

Now the scattering channel can be labeled by the eigenmodes, which could be either traveling or evanescent. The scattering matrix $\mathbf{S}(E)$ provides a complete description for the scattering from one traveling eigenmode into another traveling eigenmode. For a general lattice model, the Fisher-Lee relations allow us to construct the scattering matrix from the lattice GF, e.g., the transmission amplitude from the right-going eigenmode $|\Phi_{+, \alpha}^{(L)}\rangle$ of the left lead into the right-going eigenmode $|\Phi_{+, \beta}^{(R)}\rangle$ of the right lead is^{25,41,52,55}:

$$S_{\beta, \alpha}^{(R, L)} |_{\alpha, \beta \in \text{traveling}} = \sqrt{\frac{v_{+, \beta}^{(R)}/a_R}{v_{+, \alpha}^{(L)}/a_L}} \langle \phi_{+, \beta}^{(R)} | \mathbf{G}_{N, 1} (\mathbf{g}^{(L)})^{-1} | \Phi_{+, \alpha}^{(L)} \rangle, \quad (28)$$

while the transmission amplitude from the left-going eigenmode $|\Phi_{-, \alpha}^{(R)}\rangle$ of the right lead into the left-going eigenmode $|\Phi_{-, \beta}^{(L)}\rangle$ of the left lead is^{25,41,52,55}

$$S_{\beta, \alpha}^{(L, R)} |_{\alpha, \beta \in \text{traveling}} = \sqrt{\frac{v_{-, \beta}^{(L)}/a_L}{v_{-, \alpha}^{(R)}/a_R}} \langle \phi_{-, \beta}^{(L)} | \mathbf{G}_{1, N} (\mathbf{g}^{(R)})^{-1} | \Phi_{-, \alpha}^{(R)} \rangle, \quad (29)$$

where a_p and $\mathbf{g}^{(p)}$ are, respectively, the unit cell thickness and *free* GF of the lead p [see Eq. (35)], and $v_{s, \alpha}^{(p)}$ is the group velocity [see Eq. (21)] of the traveling eigenmode $|\Phi_{s, \alpha}^{(p)}\rangle$ in the lead p .

B. Our GF approach: Key ideas

Let us assume that there is a local excitation at the unit cell m_0 , as described by an M_{m_0} -dimensional column vector $|\Phi_{\text{loc}}\rangle_{m_0}$. This excitation generates a casual scattering wave $|\Phi(m)\rangle$ that has an energy E and obeys the Schrödinger equation with a local source at m_0 :

$$-\mathbf{H}_{m, m-1} |\Phi(m-1)\rangle + (z - \mathbf{H}_{m, m}) |\Phi(m)\rangle - \mathbf{H}_{m, m+1} |\Phi(m+1)\rangle = \delta_{m, m_0} |\Phi_{\text{loc}}\rangle_{m_0}. \quad (30)$$

The solution is given by

$$|\Phi(m)\rangle = \mathbf{G}_{m, m_0} |\Phi_{\text{loc}}\rangle_{m_0}, \quad (31)$$

e.g., for a unit excitation of the α th basis state, as described by $|\Phi_{\text{loc}}\rangle_{m_0} = [0, \dots, 1, 0, \dots, 0]^T$ (only the α th element is nonzero), Eq. (31) gives $|\Phi(m)\rangle$ as the α th column of \mathbf{G}_{m, m_0} .

Equation (31) shows that the GF can be immediately obtained once the scattering state is determined, e.g., based on physical considerations on how the local excitation evolves to the scattering state. For example, if the local excitation $|\Phi_{\text{loc}}\rangle_{m_0}$ occurs inside a lead, then it first generates an outgoing zeroth-order wave consisting of a left-going one at $m \leq m_0$ and a right-going one at $m \geq m_0$. For an infinite lead, there are no scatterers, so this outgoing wave is the total scattering state

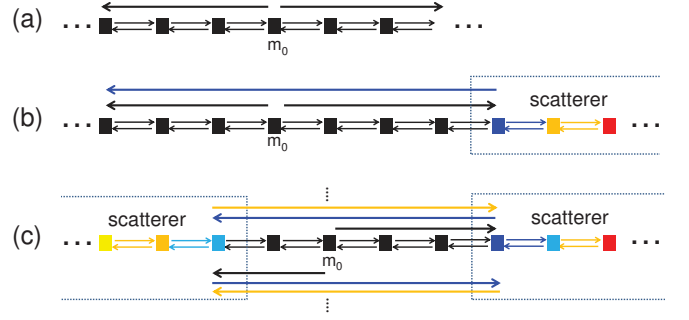


FIG. 3. Scattering state emanating from a local excitation at m_0 in an infinite lead (a), a semi-infinite lead connected to a scatterer on its right (b), and a finite lead sandwiched between two scatterers. The zeroth-order, first-order, and second-order partial waves are denoted by black, blue, and orange arrows, respectively.

[Fig. 3(a)]. For a semi-infinite lead connected to a scatterer on its right [Fig. 3(b)], the zeroth-order right-going wave will produce a first-order reflection wave, so the total scattering state in the lead is the sum of the zeroth-order and first-order waves. More generally, for a finite lead sandwiched between two scatterers [Fig. 3(c)], the right-going (left-going) zeroth-order wave will propagate to the right (left) scatterer and produce a first-order reflection wave, which in turn will propagate to the left (right) scatterer and then produce high-order reflection waves. The total scattering state would be the sum of all these waves.

In contrast to the commonly used RGF that treats the *entire* central region (regarded as a large scatterer) numerically [see Fig. 2(a) for an example], our method need only regard each *truly disordered* region as a scatterer for numerical treatment, while all the periodic subregions [such as the middle lead in Fig. 2(a)] inside the central region can be treated semianalytically by fully utilizing the translational invariance of these subregions. Physically, the wave propagation inside these periodic subregions leads to complicated multiple reflection between different scatterers, which is difficult to handle analytically in the standard RGF technique in which the GF is treated as a matrix. By contrast, in our approach, it is straightforward to perform analytically an infinite summation over all the multiple reflection waves, so that the time cost can be significantly reduced. This approach also provides a physically transparent expansion of the GF $G(E)$ in terms of a generalized scattering matrix $\mathbf{S}(E)$, which can be regarded as a reverse of the well-known Fisher-Lee relations^{14–17,25,41,55} (see Fig. 1). In the next section, we will establish the procedures for calculating the GF within this framework in a physically transparent way.

III. OUR GREEN'S FUNCTION APPROACH

Our GF approach essentially consists of two steps: generation of the zeroth-order outgoing partial wave by the local excitation and its propagation in the leads and scattering by the scatterers. The wave function mode matching approach^{52–55}

has developed useful tools to describe the latter process. Below we begin with an infinite lead, then we consider an infinite system containing a single scatterer. Finally, we give the analytical construction rule for the GF of a general layered system containing an arbitrary number of scatterers.

A. Infinite lead

Suppose that the lead is characterized by the unit cell Hamiltonian \mathbf{h} and hopping \mathbf{t} . The scattering state $|\Phi(m)\rangle$ emanating from a local excitation $|\Phi_{\text{loc}}\rangle_{m_0}$ is determined by the Schrödinger equation with a local source at m_0 :

$$-\mathbf{t}^\dagger|\Phi(m-1)\rangle + (z - \mathbf{h})|\Phi(m)\rangle - \mathbf{t}|\Phi(m+1)\rangle = \delta_{m,m_0}|\Phi_{\text{loc}}\rangle_{m_0}. \quad (32)$$

In either the left region ($m \leq m_0 - 1$) or the right region ($m \geq m_0 + 1$), the local source vanishes, so the general solution would be a linear combination of left-going and right-going eigenmodes with energy E . However, by causality considerations (due to the infinitesimal imaginary part of the energy $z = E + i0^+$), the solution in the left (right) region should be a left-going (right-going) wave [see Fig. 3(a)]:

$$|\Phi(m)\rangle|_{m \leq m_0 - 1} = (\mathbf{P}_-)^{m-m_0}|\Phi(m_0)\rangle, \quad (33)$$

$$|\Phi(m)\rangle|_{m \geq m_0 + 1} = (\mathbf{P}_+)^{m-m_0}|\Phi(m_0)\rangle. \quad (34)$$

Substituting into Eq. (32) gives $|\Phi(m_0)\rangle = \mathbf{g}|\Phi_{\text{loc}}\rangle_{m_0}$, where

$$\mathbf{g} \equiv (z - \mathbf{h} - \mathbf{t}^\dagger \mathbf{P}_-^{-1} - \mathbf{t} \mathbf{P}_+)^{-1} \quad (35a)$$

$$= [\mathbf{t}(\mathbf{P}_- - \mathbf{P}_+)]^{-1} = [\mathbf{t}^\dagger(\mathbf{P}_+^{-1} - \mathbf{P}_-^{-1})]^{-1}. \quad (35b)$$

Here we have used the equality⁵⁵

$$E - \mathbf{h} = \mathbf{t}^\dagger \mathbf{P}_\pm^{-1} + \mathbf{t} \mathbf{P}_\pm \quad (36)$$

in arriving at Eq. (35b). From the scattering wave function, we immediately identify the GF of an infinite lead as

$$\mathbf{g}_{m,m_0} \equiv \begin{cases} \mathbf{P}_+^{m-m_0} \mathbf{g} & (m \geq m_0), \\ \mathbf{P}_-^{m-m_0} \mathbf{g} & (m \leq m_0). \end{cases} \quad (37)$$

This recovers the previous result^{41,55} obtained by directly solving the equations of motion of the GF. For convenience, hereafter we call \mathbf{g}_{m,m_0} the *free GF* of the lead since it describes the generation of the zeroth-order outgoing wave $|\Phi(m)\rangle = \mathbf{g}_{m,m_0}|\Phi_{\text{loc}}\rangle_{m_0}$ from a local excitation inside this lead.

B. Single scatterer

Let us consider a scatterer C connected to two semi-infinite leads L and R [Fig. 4(a)]. The left (right) surface of the scatterer is m_L (m_R). The unit cell Hamiltonian and hopping inside the left (right) lead are \mathbf{h}_L and \mathbf{t}_L (\mathbf{h}_R and \mathbf{t}_R).

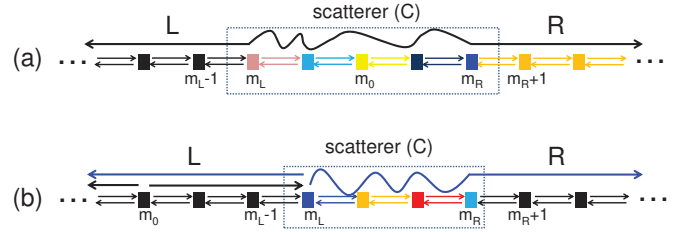


FIG. 4. Scattering state emanating from a local excitation inside a scatterer (a) or a lead (b). The black arrows denote the zeroth-order partial wave and the blue arrows denote the first-order partial wave due to scattering.

1. Local excitation inside the scatterer

The zeroth-order outgoing wave $|\Phi(m)\rangle$ emanating from a local excitation $|\Phi_{\text{loc}}\rangle_{m_0}$ at $m_0 \in C$ obeys Eq. (30) with $m_L \leq m \leq m_R$, i.e., inside the scatterer. Inside the left lead, $|\Phi(m)\rangle$ obeys

$$-\mathbf{t}_L^\dagger|\Phi(m-1)\rangle + (z - \mathbf{h}_L)|\Phi(m)\rangle - \mathbf{t}_L|\Phi(m+1)\rangle = 0 \quad (38)$$

with $m \leq m_L - 1$. Inside the right lead, $|\Phi(m)\rangle$ obeys

$$-\mathbf{t}_R^\dagger|\Phi(m-1)\rangle + (z - \mathbf{h}_R)|\Phi(m)\rangle - \mathbf{t}_R|\Phi(m+1)\rangle = 0 \quad (39)$$

with $m \geq m_R + 1$. By causality, the solution in the left (right) lead is a left-going (right-going) wave [see Fig. 4(a)]:

$$|\Phi(m)\rangle|_{m \in L} = (\mathbf{P}_-^{(L)})^{m-m_L}|\Phi(m_L)\rangle, \quad (40)$$

$$|\Phi(m)\rangle|_{m \in R} = (\mathbf{P}_+^{(R)})^{m-m_R}|\Phi(m_R)\rangle, \quad (41)$$

where $\mathbf{P}_\pm^{(p)}$ are propagation matrices of lead p [see Eq. (22)]. Substituting $|\Phi(m_L - 1)\rangle = (\mathbf{P}_-^{(L)})^{-1}|\Phi(m_L)\rangle$ and $|\Phi(m_R + 1)\rangle = \mathbf{P}_+^{(R)}|\Phi(m_R)\rangle$ into Eq. (30) gives a closed set of equations for $|\Phi(m)\rangle$ inside the scatterer. The solution is

$$|\Phi(m)\rangle|_{m \in C} = \mathbb{G}_{m,m_0}|\Phi_{\text{loc}}\rangle_{m_0}, \quad (42)$$

where

$$\mathbb{G} \equiv (z - \mathbb{H})^{-1} \quad (43)$$

and \mathbb{H} is the effective Hamiltonian for the scatterer: it is equal to the scatterer part of the system Hamiltonian \mathbf{H} , except for the two surface unit cells:

$$\mathbb{H}_{m_L,m_L} = \mathbf{H}_{m_L,m_L} + \mathbf{t}_L^\dagger(\mathbf{P}_-^{(L)})^{-1}, \quad (44a)$$

$$\mathbb{H}_{m_R,m_R} = \mathbf{H}_{m_R,m_R} + \mathbf{t}_R\mathbf{P}_+^{(R)}. \quad (44b)$$

Since \mathbb{G} converts a local excitation inside the scatterer into a scattering state inside the scatterer, we call it the *conversion matrix* of the scatterer. Comparing Eqs. (43) and (44) to Eqs. (6), (7), and (16), we see that \mathbb{G} is just the scatterer part of the GF. Actually, from the scattering wave function, we immediately identify the GF:

$$\mathbf{G}_{m \in C, m_0 \in C} = \mathbb{G}_{m,m_0}, \quad (45a)$$

$$\mathbf{G}_{m \in L, m_0 \in S} = (\mathbf{P}_-^{(L)})^{m-m_L} \mathbb{G}_{m_L,m_0}, \quad (45b)$$

$$\mathbf{G}_{m \in R, m_0 \in S} = (\mathbf{P}_+^{(R)})^{m-m_R} \mathbb{G}_{m_R,m_0}. \quad (45c)$$

Equation (45b) shows that the local excitation first evolves to an outgoing wave $\mathbb{G}_{m_L, m_0} |\Phi_{\text{loc}}\rangle_{m_0}$ at the left surface of the scatterer, and then propagates to the unit cell m as $(\mathbf{P}_-^{(L)})^{m-m_L} \mathbb{G}_{m_L, m_0} |\Phi_{\text{loc}}\rangle_{m_0}$. Equation (45c) has a similar physical interpretation.

2. Local excitation in the lead

As shown in Fig. 4(b), for $m_0 \in L$, the local excitation first generates a zeroth-order outgoing wave in the left lead: $|\Phi^{(0)}(m)\rangle = \mathbf{g}_{m, m_0}^{(L)} |\Phi_{\text{loc}}\rangle_{m_0}$, where $\mathbf{g}_{m, m_0}^{(p)}$ is the free GF of the lead p [Eq. (37)]. Next the right-going partial wave reaches the left surface of C and evolves into a scattering state $|\Psi(m)\rangle$. For $m_0 \in R$, the local excitation first generates a zeroth-order outgoing wave in the right lead: $|\Phi^{(0)}(m)\rangle = \mathbf{g}_{m, m_0}^{(R)} |\Phi_{\text{loc}}\rangle_{m_0}$. Next the left-going partial wave reaches the right surface of C and evolves to a scattering state $|\Psi(m)\rangle$. For either case, the total scattering state $|\Phi(m)\rangle$ emanating from the local excitation is the sum of the unscattered zeroth-order partial wave and the scattering state $|\Psi(m)\rangle$. The central issue is to determine the scattering state emanating from a known incident wave, in a way similar to the wave function mode matching approach to mesoscopic quantum transport^{52–55}.

First we consider the scattering state $|\Psi(m)\rangle$ emanating from a right-going incident wave $|\Phi_{\text{in}}(m)\rangle$ in the left lead. The key observation is that for arbitrary $m_1 \leq m_L$, the local excitation $|\Phi_{\text{loc}}\rangle_{m_1} \equiv (\mathbf{g}^{(L)})^{-1} |\Phi_{\text{in}}(m_1)\rangle$ generates a right-going partial wave $|\tilde{\Phi}(m)\rangle_{m \geq m_1} = (\mathbf{P}_+^{(L)})^{m-m_1} |\Phi_{\text{in}}(m_1)\rangle$ that is equal to $|\Phi_{\text{in}}(m)\rangle_{m \geq m_1}$. Therefore, in the region $m \geq m_1$, the scattering state emanating from $|\Phi_{\text{in}}(m)\rangle$ is the same as the scattering state emanating from this local excitation (see Appendix B for a rigorous proof). Taking $m_1 = m_L$ immediately gives

$$|\Psi(m)\rangle_{m \in C} = \mathbb{G}_{m, m_L} (\mathbf{g}^{(L)})^{-1} |\Phi_{\text{in}}(m_L)\rangle, \quad (46)$$

i.e., first the incident wave amplitude $|\Phi_{\text{in}}(m_L)\rangle$ is converted back to a local excitation $|\Phi_{\text{loc}}\rangle_{m_L} \equiv (\mathbf{g}^{(L)})^{-1} |\Phi_{\text{in}}(m_L)\rangle$, then the conversion matrix \mathbb{G} of the scatterer further converts it to the total scattering state $|\Psi(m)\rangle_{m \in C}$ according to Eq. (42). Inside the left lead, $|\Psi(m)\rangle$ is the sum of the right-going incident wave and a left-going reflection partial wave $|\Phi_r(m)\rangle_{m \in L} = (\mathbf{P}_-^{(L)})^{m-m_L} |\Phi_r(m_L)\rangle$, where

$$\begin{aligned} |\Phi_r(m_L)\rangle &= |\Psi(m_L)\rangle - |\Phi_{\text{in}}(m_L)\rangle \\ &= [\mathbb{G}_{m_L, m_L} (\mathbf{g}^{(L)})^{-1} - \mathbf{I}] |\Phi_{\text{in}}(m_L)\rangle. \end{aligned} \quad (47)$$

Inside the right lead, $|\Psi(m)\rangle$ is the right-going transmission wave: $|\Psi(m)\rangle_{m \in R} = (\mathbf{P}_+^{(R)})^{m-m_R} |\Psi(m_R)\rangle$, where

$$|\Psi(m_R)\rangle = \mathbb{G}_{m_R, m_L} (\mathbf{g}^{(L)})^{-1} |\Phi_{\text{in}}(m_L)\rangle. \quad (48)$$

Similarly, we can derive the scattering state $|\Psi(m)\rangle$ emanating from a left-going incident wave $|\Phi_{\text{in}}(m)\rangle$ in the right lead. Inside the scatterer, the scattering state is

$$|\Psi(m)\rangle_{m \in S} = \mathbb{G}_{m, m_R} (\mathbf{g}^{(R)})^{-1} |\Phi_{\text{in}}(m_R)\rangle, \quad (49)$$

as if it emanated from a local excitation $|\Phi_{\text{loc}}\rangle_{m_R} \equiv (\mathbf{g}^{(R)})^{-1} |\Phi_{\text{in}}(m_R)\rangle$ at the right surface of the scatterer [cf. Eq.

(42)]. Inside the right lead, $|\Psi(m)\rangle$ is the sum of the left-going incident wave and a right-going reflection partial wave $|\Phi_r(m)\rangle_{m \in R} = (\mathbf{P}_+^{(R)})^{m-m_R} |\Phi_r(m_R)\rangle$, where

$$\begin{aligned} |\Phi_r(m_R)\rangle &= |\Psi(m_R)\rangle - |\Phi_{\text{in}}(m_R)\rangle \\ &= [\mathbb{G}_{m_R, m_R} (\mathbf{g}^{(R)})^{-1} - \mathbf{I}] |\Phi_{\text{in}}(m_R)\rangle. \end{aligned} \quad (50)$$

Inside the left lead, $|\Psi(m)\rangle$ is the left-going transmission wave: $|\Psi(m)\rangle = (\mathbf{P}_-^{(L)})^{m-m_L} |\Psi(m_L)\rangle$, where

$$|\Psi(m_L)\rangle = \mathbb{G}_{m_L, m_R} (\mathbf{g}^{(R)})^{-1} |\Phi_{\text{in}}(m_R)\rangle. \quad (51)$$

Using the above results, the scattering state emanating from the zeroth-order right-going partial wave in the left lead is given by Eqs. (46)-(48) with $|\Phi_{\text{in}}(m_L)\rangle \equiv |\Phi^{(0)}(m_L)\rangle$. This allows us to identify the GF

$$\mathbf{G}_{m \in C, m_0 \in L} = \mathbb{G}_{m, m_L} (\mathbf{g}^{(L)})^{-1} \mathbf{g}_{m_L, m_0}^{(L)}, \quad (52a)$$

$$\mathbf{G}_{m \in R, m_0 \in L} = (\mathbf{P}_+^{(R)})^{m-m_R} \mathbb{G}_{m_R, m_L} (\mathbf{g}^{(L)})^{-1} \mathbf{g}_{m_L, m_0}^{(L)}, \quad (52b)$$

$$\mathbf{G}_{m \in L, m_0 \in L} = \mathbf{g}_{m, m_0}^{(L)} + (\mathbf{P}_-^{(L)})^{m-m_L} [\mathbb{G}_{m_L, m_L} (\mathbf{g}^{(L)})^{-1} - \mathbf{I}] \mathbf{g}_{m_L, m_0}^{(L)}. \quad (52c)$$

These expressions have clear physical interpretations; e.g., Eq. (52a) shows that the local excitation $|\Phi_{\text{loc}}\rangle_{m_0 \in L}$ first evolves to a right-going partial wave and propagates rightward to the left surface of the scatterer as $\mathbf{g}_{m_L, m_0}^{(L)} |\Phi_{\text{loc}}\rangle_{m_0 \in L}$. There it is converted back to a local excitation by $(\mathbf{g}^{(L)})^{-1}$, and finally the conversion matrix of the scatterer \mathbb{G}_{m, m_L} further converts it to the scattering state inside the scatterer. As another example, Eq. (52c) shows that the total scattering wave inside the left lead is the sum of the zeroth-order partial wave $\mathbf{g}_{m, m_0}^{(L)} |\Phi_{\text{loc}}\rangle_{m_0}$ and the reflection partial wave: first the local excitation $|\Phi_{\text{loc}}\rangle_{m_0 \in L}$ evolves to a zeroth-order partial wave and then propagates rightwards to the left surface of the scatterer as $\mathbf{g}_{m_L, m_0}^{(L)} |\Phi_{\text{loc}}\rangle_{m_0 \in L}$, then $\mathbb{G}_{m_L, m_L} (\mathbf{g}^{(L)})^{-1} - \mathbf{I}$ converts it to the reflection wave. Finally, $(\mathbf{P}_-^{(L)})^{m-m_L}$ propagates this reflection wave leftward to m .

Similarly, the scattering state emanating from the zeroth-order left-going partial wave in the right lead is given by Eqs. (49)-(51) with $|\Phi_{\text{in}}(m_R)\rangle \equiv |\Phi^{(0)}(m_R)\rangle$. This allows us to identify the GF

$$\mathbf{G}_{m \in C, m_0 \in R} = \mathbb{G}_{m, m_R} (\mathbf{g}^{(R)})^{-1} \mathbf{g}_{m_R, m_0}^{(R)}, \quad (53a)$$

$$\mathbf{G}_{m \in L, m_0 \in R} = (\mathbf{P}_-^{(L)})^{m-m_L} \mathbb{G}_{m_L, m_R} (\mathbf{g}^{(R)})^{-1} \mathbf{g}_{m_R, m_0}^{(R)}, \quad (53b)$$

$$\mathbf{G}_{m \in R, m_0 \in R} = \mathbf{g}_{m, m_0}^{(R)} + (\mathbf{P}_+^{(R)})^{m-m_R} [\mathbb{G}_{m_R, m_R} (\mathbf{g}^{(R)})^{-1} - \mathbf{I}] \mathbf{g}_{m_R, m_0}^{(R)}. \quad (53c)$$

These can be interpreted in a similar way to Eqs. (45) and (52).

The above results cover previous results as special cases. For example, by directly solving the equation of motion, Sanvito *et al.*⁴¹ and Krstić *et al.*⁷⁷ obtain the GF of an infinite lead [Eq. (37)] and a semi-infinite lead consisting of the unit cells $m \leq 0$ ($m \geq 0$) [Eq. (C2)], which can be regarded as a single-unit-cell scatterer at $m = 0$ connected to a semi-infinite left (right) lead. Khomyakov *et al.*⁵⁵ further obtained the GF

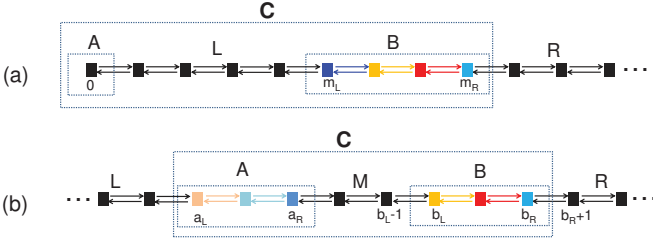


FIG. 5. (a) A scatterer B connected to a finite left lead L and a semi-infinite right lead R can be regarded as a composite scatterer C connected to one semi-infinite right lead R . (b) Two scatterers sandwiched between three leads L, M, R can be regarded as a composite scatterer C connected to two semi-infinite leads L and R .

across a single scatterer [Eq. (52b)]. A sharp interface between a semi-infinite left lead and a semi-infinite right lead can also be regarded as a single-unit-cell scatterer connected to two semi-infinite leads. For reference, the explicit expressions of the GFs for these simple cases are given in Appendix C.

C. Multiple scatterers

A general layered system containing an arbitrary number of scatterers can be regarded as a *composite* scatterer connected to one or two semi-infinite leads, e.g., a scatterer B connected to a finite lead L and a semi-infinite lead R can be regarded as a composite scatterer $C = (A + B)$ connected to one semi-infinite right lead [Fig. 5(a)], while two scatterers sandwiched between three leads can be regarded as a composite scatterer $C = (A + B)$ connected to two semi-infinite leads [Fig. 5(b)]. Therefore, we can use Eqs. (45), (52), and (53) to obtain the GF of the entire system once the conversion matrix of the composite scatterer is known. In the RGF method, the conversion matrix (which coincides with the GF of the infinite system within the composite scatterer) is calculated by a numerical iteration algorithm that builds up the composite scatterer slice by slice, thus the time cost increases linearly with the total length of the composite scatterer. Here the physical transparency of our approach allows us to treat the multiple reflection between different scatterers analytically, so that the conversion matrix of a composite scatterer can be obtained by combining the conversion matrices of the constituent scatterers analytically with a significantly reduced time cost. The basic step is the combination of the conversion matrices $\mathbb{G}^{(A)}$ and $\mathbb{G}^{(B)}$ of two scatterers A and B into the conversion matrix \mathbb{G} of a composite scatterer $C \equiv (A + B)$.

1. Combining conversion matrices

As shown in Fig. 5(b), the left and right surfaces of the scatterer A (B) are a_L and a_R (b_L and b_R) and the three leads sandwiching the scatterers are the semi-infinite left lead L , the middle lead M , and the semi-infinite right lead R . For a local excitation $|\Phi_{\text{loc}}\rangle_{m_0}$ at $m_0 \in C$, the total scattering state inside C

is $|\Psi(m)\rangle|_{m \in C} = \mathbb{G}_{m,m_0}|\Phi_{\text{loc}}\rangle_{m_0}$, which allows us to identify \mathbb{G} once the total scattering state has been obtained. For $m_0 \in A$, the local excitation first produces a zeroth-order partial wave in A and M . Next the right-going partial wave in M undergoes multiple reflections between B and A and finally evolves to a scattering state. For $m_0 \in B$, the local excitation first produces a zeroth-order partial wave inside B and M . Next the left-going partial wave in M undergoes multiple reflections and finally evolves to a scattering state. For $m_0 \in M$, the local excitation first produces a zeroth-order outgoing partial wave in M . Next the left- and right-going partial waves each undergo multiple reflections and evolve to a scattering state. For each case, the total scattering state emanating from the local excitation is the sum of the unscattered zeroth-order partial wave and the scattering state(s) emanating from the scattered zeroth-order partial wave. Therefore, the key issue is to calculate the scattering state $|\Psi(m)\rangle$ emanating from a right- or left-going incident wave $|\Phi_{\text{in}}(m)\rangle$ in the middle lead through multiple reflections between A and B .

In the middle lead, $|\Psi(m)\rangle$ is the sum of the right-going part $|\Psi_+(m)\rangle = (\mathbf{P}_+^{(M)})^{m-b_L}|\Psi_+(b_L)\rangle$ and the left-going part $|\Psi_-(m)\rangle = (\mathbf{P}_-^{(M)})^{m-a_R}|\Psi_-(a_R)\rangle$. Inside the scatterer A , $|\Psi(m)\rangle$ is equal to the scattering state emanating from the total incident wave $|\Psi_-(a_R)\rangle$ on A [cf. Eq. (49)]:

$$|\Psi(m)\rangle|_{m \in A} = \mathbb{G}_{m,a_R}^{(A)}(\mathbf{g}^{(M)})^{-1}|\Psi_-(a_R)\rangle, \quad (54)$$

Inside the scatterer B , $|\Psi(m)\rangle$ is equal to the scattering state emanating from the total incident wave $|\Psi_+(b_L)\rangle$ on B [cf. Eq. (46)]:

$$|\Psi(m)\rangle|_{m \in B} = \mathbb{G}_{m,b_L}^{(B)}(\mathbf{g}^{(M)})^{-1}|\Psi_+(b_L)\rangle. \quad (55)$$

Therefore, the scattering state inside C is completely determined by $|\Psi_+(b_L)\rangle$ and $|\Psi_-(a_R)\rangle$. For brevity, we introduce the reflection matrices

$$\mathbb{R}_B \equiv \mathbb{G}_{b_L,b_L}^{(B)}(\mathbf{g}^{(M)})^{-1} - \mathbf{I}, \quad (56)$$

$$\mathbb{R}_A \equiv \mathbb{G}_{a_R,a_R}^{(A)}(\mathbf{g}^{(M)})^{-1} - \mathbf{I}. \quad (57)$$

The former (latter) converts a right-going (left-going) incident wave on the left (right) surface of B (A) to a left-going (right-going) reflection wave. To describe the multiple reflections between A and B , we introduce the following renormalized propagation matrices that incorporate multiple reflections:

$$\mathbf{P}_{b_L \leftarrow a_R} \equiv [1 - (\mathbf{P}_+^{(M)})^{\Delta m} \mathbb{R}_A (\mathbf{P}_-^{(M)})^{-\Delta m} \mathbb{R}_B]^{-1} (\mathbf{P}_+^{(M)})^{\Delta m}, \quad (58)$$

$$\mathbf{P}_{a_R \leftarrow b_L} \equiv [1 - (\mathbf{P}_-^{(M)})^{-\Delta m} \mathbb{R}_B (\mathbf{P}_+^{(M)})^{\Delta m} \mathbb{R}_A]^{-1} (\mathbf{P}_-^{(M)})^{-\Delta m}, \quad (59)$$

$$\mathbf{P}_{a_R \leftarrow a_R} \equiv (\mathbf{P}_-^{(M)})^{-\Delta m} \mathbb{R}_B \mathbf{P}_{b_L \leftarrow a_R} = \mathbf{P}_{a_R \leftarrow b_L} \mathbb{R}_B (\mathbf{P}_+^{(M)})^{\Delta m}, \quad (60)$$

$$\mathbf{P}_{b_L \leftarrow b_L} \equiv (\mathbf{P}_+^{(M)})^{\Delta m} \mathbb{R}_A \mathbf{P}_{a_R \leftarrow b_L} = \mathbf{P}_{b_L \leftarrow a_R} \mathbb{R}_A (\mathbf{P}_-^{(M)})^{-\Delta m}, \quad (61)$$

where $\Delta m \equiv b_L - a_R$ is the distance between A and B . For example, the renormalized propagation matrix $\mathbf{P}_{b_L \leftarrow a_R}$ from a_R to b_L is the sum of the free propagation term $(\mathbf{P}_+^{(M)})^{\Delta m}$, the propagation term with two reflections $(\mathbf{P}_+^{(M)})^{\Delta m} \mathbb{R}_A (\mathbf{P}_-^{(M)})^{-\Delta m} \mathbb{R}_B (\mathbf{P}_+^{(M)})^{\Delta m}$, and so on.

Using the above notations, when $|\Phi_{\text{in}}(m)\rangle$ is a right-going incident wave on B , we have

$$|\Psi_+(b_L)\rangle = (1 + \mathbf{P}_{b_L \leftarrow b_L} \mathbb{R}_B) |\Phi_{\text{in}}(b_L)\rangle, \quad (62)$$

$$|\Psi_-(a_R)\rangle = \mathbf{P}_{a_R \leftarrow b_L} \mathbb{R}_B |\Phi_{\text{in}}(b_L)\rangle. \quad (63)$$

When $|\Phi_{\text{in}}(m)\rangle$ is a left-going incident wave on A , we have

$$|\Psi_+(b_L)\rangle = \mathbf{P}_{b_L \leftarrow a_R} \mathbb{R}_A |\Phi_{\text{in}}(a_R)\rangle, \quad (64)$$

$$|\Psi_-(a_R)\rangle = (1 + \mathbf{P}_{a_R \leftarrow a_R} \mathbb{R}_A) |\Phi_{\text{in}}(a_R)\rangle. \quad (65)$$

Using the above results together with $|\Phi(m)\rangle|_{m \in C} = \mathbb{G}_{m,m_0} |\Phi_{\text{loc}}\rangle_{m_0}$, we identify

$$\mathbb{G}_{m \in B, m_0 \in A} = \mathbb{G}_{m, b_L}^{(B)} (\mathbf{g}^{(M)})^{-1} \mathbf{P}_{b_L \leftarrow a_R} \mathbb{G}_{a_R, m_0}^{(A)}, \quad (66a)$$

$$\mathbb{G}_{m \in A, m_0 \in B} = \mathbb{G}_{m, a_R}^{(A)} (\mathbf{g}^{(M)})^{-1} \mathbf{P}_{a_R \leftarrow b_L} \mathbb{G}_{b_L, m_0}^{(B)}, \quad (66b)$$

$$\mathbb{G}_{m \in A, m_0 \in A} = \mathbb{G}_{m, m_0}^{(A)} + \mathbb{G}_{m, a_R}^{(A)} (\mathbf{g}^{(M)})^{-1} \mathbf{P}_{a_R \leftarrow a_R} \mathbb{G}_{a_R, m_0}^{(A)}, \quad (66c)$$

$$\mathbb{G}_{m \in B, m_0 \in B} = \mathbb{G}_{m, m_0}^{(B)} + \mathbb{G}_{m, b_L}^{(B)} (\mathbf{g}^{(M)})^{-1} \mathbf{P}_{b_L \leftarrow b_L} \mathbb{G}_{b_L, m_0}^{(B)}, \quad (66d)$$

$$\mathbb{G}_{m \in M, m_0 \in A} = [(\mathbf{P}_+^{(M)})^{m-a_R} (1 + \mathbb{R}_A \mathbf{P}_{a_R \leftarrow a_R}) + (\mathbf{P}_-^{(M)})^{m-b_L} \mathbb{R}_B \mathbf{P}_{b_L \leftarrow a_R}] \mathbb{G}_{a_R, m_0}^{(A)}, \quad (66e)$$

$$\mathbb{G}_{m \in A, m_0 \in M} = \mathbb{G}_{m, a_R}^{(A)} (\mathbf{g}^{(M)})^{-1} [(1 + \mathbf{P}_{a_R \leftarrow a_R} \mathbb{R}_A) \mathbf{g}_{a_R, m_0}^{(M)} + \mathbf{P}_{a_R \leftarrow b_L} \mathbb{R}_B \mathbf{g}_{b_L, m_0}^{(M)}], \quad (66f)$$

$$\mathbb{G}_{m \in M, m_0 \in B} = [(\mathbf{P}_-^{(M)})^{m-b_L} (1 + \mathbb{R}_B \mathbf{P}_{b_L \leftarrow b_L}) + (\mathbf{P}_+^{(M)})^{m-a_R} \mathbb{R}_A \mathbf{P}_{a_R \leftarrow b_L}] \mathbb{G}_{b_L, m_0}^{(B)}, \quad (66g)$$

$$\mathbb{G}_{m \in B, m_0 \in M} = \mathbb{G}_{m, b_L}^{(B)} (\mathbf{g}^{(M)})^{-1} [(1 + \mathbf{P}_{b_L \leftarrow b_L} \mathbb{R}_B) \mathbf{g}_{b_L, m_0}^{(M)} + \mathbf{P}_{b_L \leftarrow a_R} \mathbb{R}_A \mathbf{g}_{a_R, m_0}^{(M)}], \quad (66h)$$

$$\begin{aligned} \mathbb{G}_{m \in M, m_0 \in M} = & \mathbf{g}_{m, m_0}^{(M)} + (\mathbf{P}_+^{(M)})^{m-a_R} \mathbb{R}_A \mathbf{P}_{a_R \leftarrow b_L} \mathbb{R}_B \mathbf{g}_{b_L, m_0}^{(M)} + (\mathbf{P}_-^{(M)})^{m-b_L} \mathbb{R}_B \mathbf{P}_{b_L \leftarrow a_R} \mathbb{R}_A \mathbf{g}_{a_R, m_0}^{(M)} \\ & + (\mathbf{P}_-^{(M)})^{m-b_L} (\mathbb{R}_B + \mathbb{R}_B \mathbf{P}_{b_L \leftarrow b_L} \mathbb{R}_B) \mathbf{g}_{b_L, m_0}^{(M)} + (\mathbf{P}_+^{(M)})^{m-a_R} (\mathbb{R}_A + \mathbb{R}_A \mathbf{P}_{a_R \leftarrow a_R} \mathbb{R}_A) \mathbf{g}_{a_R, m_0}^{(M)}. \end{aligned} \quad (66i)$$

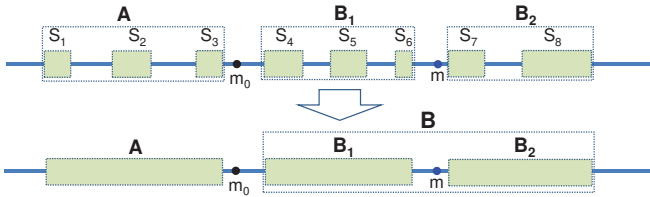


FIG. 6. Green's function \mathbb{G}_{m,m_0} of an infinite system containing eight scatterers, where m_0 and m are both inside the leads or at the surfaces of the scatterers.

These equations can be interpreted in a physically transparent way. For example, Eq. (66a) shows that the local excitation at $m_0 \in A$ evolves to the scattering wave at $m \in B$ through the following steps: first it is converted by $\mathbb{G}_{a_R, m_0}^{(A)}$ to a zeroth-order partial wave at a_R , next it undergoes renormalized propagation from a_R to b_L , and finally it is converted back to a local excitation and then to the scattering wave at $m \in B$.

2. Analytical construction rule for multiple scatterers

By repeatedly using Eq. (66), the conversion matrix of a composite scatterer can be obtained analytically as functions of the conversion matrices of the constituent scatterers. In

particular, the four surface elements of the conversion matrix of the composite scatterer can be immediately obtained from those of the constituent scatterers:

$$\mathbb{G}_{b_R, a_L} = \mathbb{G}_{b_R, b_L}^{(B)} (\mathbf{g}^{(M)})^{-1} \mathbf{P}_{b_L \leftarrow a_R} \mathbb{G}_{a_R, a_L}^{(A)}, \quad (67a)$$

$$\mathbb{G}_{a_L, b_R} = \mathbb{G}_{a_L, a_R}^{(A)} (\mathbf{g}^{(M)})^{-1} \mathbf{P}_{a_R \leftarrow b_L} \mathbb{G}_{b_L, b_R}^{(B)}, \quad (67b)$$

$$\mathbb{G}_{a_L, a_L} = \mathbb{G}_{a_L, a_L}^{(A)} + \mathbb{G}_{a_L, a_R}^{(A)} (\mathbf{g}^{(M)})^{-1} \mathbf{P}_{a_R \leftarrow a_R} \mathbb{G}_{a_R, a_L}^{(A)}, \quad (67c)$$

$$\mathbb{G}_{b_R, b_R} = \mathbb{G}_{b_R, b_R}^{(B)} + \mathbb{G}_{b_R, b_L}^{(B)} (\mathbf{g}^{(M)})^{-1} \mathbf{P}_{b_L \leftarrow b_L} \mathbb{G}_{b_L, b_R}^{(B)}, \quad (67d)$$

while the latter can be calculated through recursive techniques [Eqs. (11) and (12)].

For example, let us consider an infinite layered system with eight scatterers S_1, S_2, \dots, S_8 and calculate its GF \mathbb{G}_{m,m_0} . For simplicity we assume that m_0 and m are both inside the leads or at the surfaces of the scatterers (see Fig. 6), so that \mathbb{G}_{m,m_0} is just the (m, m_0) element of the conversion matrix \mathbb{G} of the composite scatterer $(S_1 + S_2 + \dots + S_8)$, i.e., $\mathbb{G}_{m,m_0} = \mathbb{G}_{m,m_0}$, and \mathbb{G}_{m,m_0} is completely determined by the surface elements of $\mathbb{G}^{(S_1)}, \dots, \mathbb{G}^{(S_8)}$, which are readily obtained through recursive techniques. First, we use Eq. (67) to calculate the surface elements of $\mathbb{G}^{(A)}$, $\mathbb{G}^{(B_1)}$, and $\mathbb{G}^{(B_2)}$ for the three composite scatterers $A \equiv (S_1 + S_2 + S_3)$, $B_1 \equiv (S_4 + S_5 + S_6)$, and $B_2 \equiv (S_7 + S_8)$. Next we regard $(B_1 + B_2)$ as a composite scatterer B and calculate $\mathbb{G}_{m, b_L}^{(B)}$ (b_L is the left surface of B) from Eq. (66e). Now the entire system contains two scatterers A and B ; thus \mathbb{G}_{m,m_0} can be obtained from Eq. (66h).

IV. INVERSE OF FISHER-LEE RELATION

In the previous section, we have developed a physically transparent and numerically efficient way to calculate the GF of a general layered system. There the GF is expressed as a matrix, i.e., in terms of the propagation matrices \mathbf{P}_\pm and conversion matrix \mathbf{g} of the leads and the conversion matrices \mathbb{G} of the scatterers. In this section, we give further physical insight into our GF approach by expressing the GF analytically in terms of a generalized scattering matrix, which describes the scattering of both traveling states and evanescent states. This could be regarded as the inverse of the well-known Fisher-Lee relations^{14–17,25,41,55} (see Fig. 1). The key is to express the conversion matrix \mathbb{G} in terms of a generalized scattering matrix.

A. Generalized scattering matrix

Let us consider an infinite system consisting of a single scatterer (with the left surface at m_L and the right surface at m_R) connected to two semi-infinite leads L and R (see Fig. 4). For a right-going eigenmode $|\Phi_{\text{in}}(m)\rangle = e^{ik_{+,a}^{(L)}(m-m_L)a_L}|\Phi_{+,a}^{(L)}\rangle$ incident on the scatterer from the left lead, the resulting scattering state $|\Psi(m)\rangle$ follows from Eqs. (46)-(48). Using Eq. (26) for the propagation matrix, we obtain

$$|\Psi(m)\rangle|_{m \geq m_R} = \sum_{\beta} e^{ik_{+,a}^{(R)}(m-m_R)a_R}|\Phi_{+,a}^{(R)}\rangle S_{\beta,\alpha}^{(RL)} \quad (68)$$

$$|\Psi(m)\rangle|_{m \leq m_L} = |\Phi_{\text{in}}(m)\rangle + \sum_{\beta} e^{ik_{-,a}^{(L)}(m-m_L)a_L}|\Phi_{-,a}^{(L)}\rangle S_{\beta,\alpha}^{(LL)}, \quad (69)$$

where

$$S_{\beta,\alpha}^{(RL)} \equiv \langle \phi_{+,a}^{(R)} | \mathbb{G}_{m_R,m_L}(\mathbf{g}^{(L)})^{-1} | \Phi_{+,a}^{(L)} \rangle \quad (70)$$

is a generalized transmission amplitude from $|\Phi_{+,a}^{(L)}\rangle$ at the left surface of the scatterer into $|\Phi_{+,a}^{(R)}\rangle$ at the right surface of the scatterer, and

$$S_{\beta,\alpha}^{(LL)} \equiv \langle \phi_{-,a}^{(L)} | [\mathbb{G}_{m_L,m_L}(\mathbf{g}^{(L)})^{-1} - \mathbf{I}] | \Phi_{+,a}^{(L)} \rangle \quad (71)$$

is a generalized reflection amplitude from $|\Phi_{+,a}^{(L)}\rangle$ into $|\Phi_{-,a}^{(L)}\rangle$ at the left surface of the scatterer. Similarly, for a left-going incident wave $|\Phi_{\text{in}}(m)\rangle = e^{ik_{-,a}^{(R)}(m-m_R)a_R}|\Phi_{-,a}^{(R)}\rangle$ in the right lead, the resulting scattering state $|\Psi(m)\rangle$ follows from Eqs. (49)-(51) as

$$|\Psi(m)\rangle|_{m \leq m_L} = \sum_{\beta} e^{ik_{-,a}^{(L)}(m-m_L)a_L}|\Phi_{-,a}^{(L)}\rangle S_{\beta,\alpha}^{(LR)}, \quad (72)$$

$$|\Psi(m)\rangle|_{m \geq m_R} = |\Phi_{\text{in}}(m)\rangle + \sum_{\beta} e^{ik_{+,a}^{(R)}(m-m_R)a_R} S_{\beta,\alpha}^{(RR)} |\Phi_{+,a}^{(R)}\rangle, \quad (73)$$

where

$$S_{\beta,\alpha}^{(LR)} \equiv \langle \phi_{-,a}^{(L)} | \mathbb{G}_{m_L,m_R}(\mathbf{g}^{(R)})^{-1} | \Phi_{-,a}^{(R)} \rangle \quad (74)$$

is a generalized transmission amplitude from $|\Phi_{-,a}^{(R)}\rangle$ at the right surface of the scatterer into $|\Phi_{-,a}^{(L)}\rangle$ at the left surface of the scatterer, and

$$S_{\beta,\alpha}^{(RR)} \equiv \langle \phi_{+,a}^{(R)} | [\mathbb{G}_{m_R,m_R}(\mathbf{g}^{(R)})^{-1} - \mathbf{I}] | \Phi_{-,a}^{(R)} \rangle \quad (75)$$

is a generalized reflection amplitude from $|\Phi_{-,a}^{(R)}\rangle$ into $|\Phi_{+,a}^{(R)}\rangle$ at the right surface of the scatterer.

Equations (70)-(75) define a generalized scattering matrix $\mathcal{S}(E)$ and express it in terms of the surface elements of the conversion matrix. They were first derived in the wave function mode matching approach^{52–54} and its connection to the GF approach was established later⁵⁵. They are valid for both traveling modes and evanescent modes. In our GF approach, these expressions have clear physical meanings. Taking $S_{\beta,\alpha}^{(RL)}$ as an example, $\mathbb{G}_{m_R,m_L}(\mathbf{g}^{(L)})^{-1}$ converts the incident eigenmode $|\Phi_{+,a}^{(L)}\rangle$ at the left surface of the scatterer back to a local excitation and then to the scattering wave at the right surface of the scatterer. Then the dual vector $\langle \phi_{+,a}^{(R)} |$ projects the scattering wave onto the eigenmode $|\Phi_{+,a}^{(R)}\rangle$ [see Eqs. (24) and (25)]. The transmission and reflection amplitudes of the unitary scattering matrix connecting two traveling eigenmodes are obtained by normalizing with respect to the current⁵⁵:

$$S_{\beta,\alpha}^{(q,p)}|_{\alpha,\beta \in \text{traveling}} = \sqrt{\frac{|v_{\text{out},\beta}^{(q)}|/a_q}{|v_{\text{in},\alpha}^{(p)}|/q_p}} S_{\beta,\alpha}^{(q,p)}, \quad (76)$$

where $s_{\text{in}} = s_{\text{out}} = +$ for $(q, p) = (R, L)$; $s_{\text{in}} = s_{\text{out}} = -$ for $(q, p) = (L, R)$; $s_{\text{in}} = +$, $s_{\text{out}} = -$ for $(q, p) = (L, L)$; and $s_{\text{in}} = -$, $s_{\text{out}} = +$ for $(q, p) = (R, R)$, thus Eqs. (70) and (74) lead to Eqs. (28) and (29), respectively.

B. Inverse of Fisher-Lee relations

The inverse of Eqs. (70)-(75) gives the surface elements of $\mathbb{G}(E)$ in terms of the generalized scattering matrix $\mathcal{S}(E)$:

$$\mathbb{G}_{m_R,m_L} = \sum_{\alpha} \left(\sum_{\beta} S_{\beta,\alpha}^{(RL)} |\Phi_{+,a}^{(R)}\rangle \right) \langle \phi_{+,a}^{(L)} | \mathbf{g}^{(L)}, \quad (77a)$$

$$\mathbb{G}_{m_L,m_R} = \sum_{\alpha} \left(\sum_{\beta} S_{\beta,\alpha}^{(LR)} |\Phi_{-,a}^{(L)}\rangle \right) \langle \phi_{-,a}^{(R)} | \mathbf{g}^{(R)}, \quad (77b)$$

$$\mathbb{G}_{m_L,m_L} = \sum_{\alpha} \left(|\Phi_{+,a}^{(L)}\rangle + \sum_{\beta} S_{\beta,\alpha}^{(LL)} |\Phi_{-,a}^{(L)}\rangle \right) \langle \phi_{+,a}^{(L)} | \mathbf{g}^{(L)}, \quad (77c)$$

$$\mathbb{G}_{m_R,m_R} = \sum_{\alpha} \left(|\Phi_{-,a}^{(R)}\rangle + \sum_{\beta} S_{\beta,\alpha}^{(RR)} |\Phi_{+,a}^{(R)}\rangle \right) \langle \phi_{-,a}^{(R)} | \mathbf{g}^{(R)}. \quad (77d)$$

These expressions have very clear physical interpretations. For example, Eq. (77a) shows that a local excitation $|\Phi_{\text{loc}}\rangle_{m_L}$ at the left surface of the scatterer evolves to the scattering wave at the right surface of the scatterer through two steps. First, it evolves to a partial wave $\mathbf{g}^{(L)}|\Phi_{\text{loc}}\rangle_{m_L}$, then

it is expanded into linear combinations of right-going eigenmodes $\sum_{\alpha} |\Phi_{+, \alpha}^{(L)}\rangle \langle \phi_{+, \alpha}^{(L)} | \mathbf{g}^{(L)} | \Phi_{\text{loc}} \rangle_{m_L}$ and each eigenmode transmits across the scatterer to its right surface as

$$|\Phi_{+, \alpha}^{(L)}\rangle \rightarrow \sum_{\beta} \mathcal{S}_{\beta, \alpha}^{(RL)} |\Phi_{+, \beta}^{(R)}\rangle. \quad (78)$$

Similarly, Eq. (77c) shows that a local excitation at the left surface of the scatterer first evolves to a partial wave $\mathbf{g}^{(L)} | \Phi_{\text{loc}} \rangle_{m_L}$, then it is expanded as $\sum_{\alpha} |\Phi_{+, \alpha}^{(L)}\rangle \langle \phi_{+, \alpha}^{(L)} | \mathbf{g}^{(L)} | \Phi_{\text{loc}} \rangle_{m_L}$ and each eigenmode evolves to a scattering wave:

$$|\Phi_{+, \alpha}^{(L)}\rangle \rightarrow |\Phi_{+, \alpha}^{(L)}\rangle + \sum_{\beta} |\Phi_{-, \beta}^{(L)}\rangle \mathcal{S}_{\beta, \alpha}^{(LL)}, \quad (79)$$

which consists of the incident wave and the reflection wave.

Next, we can express other blocks of the GF, i.e., Eqs. (45)-(53) with m, m_0 inside the leads, in terms of the generalized scattering matrix:

$$\begin{aligned} \mathbf{G}_{m \in R, m_0 \in L} &= \sum_{\alpha} \left(\sum_{\beta} e^{ik_{+, \beta}^{(R)}(m-m_R)a} |\Phi_{+, \beta}^{(R)}\rangle \mathcal{S}_{\beta, \alpha}^{(RL)} \right) \\ &\times e^{ik_{+, \alpha}^{(L)}(m_L-m_0)a} \langle \phi_{+, \alpha}^{(L)} | \mathbf{g}^{(L)}, \end{aligned} \quad (80a)$$

$$\begin{aligned} \mathbf{G}_{m \in L, m_0 \in R} &= \sum_{\alpha} \left(\sum_{\beta} e^{ik_{-, \beta}^{(L)}(m-m_L)a} |\Phi_{-, \beta}^{(L)}\rangle \mathcal{S}_{\beta, \alpha}^{(LR)} \right) \\ &\times e^{ik_{-, \alpha}^{(R)}(m_R-m_0)a} \langle \phi_{-, \alpha}^{(R)} | \mathbf{g}^{(R)}, \end{aligned} \quad (80b)$$

$$\begin{aligned} \mathbf{G}_{m \in L, m_0 \in L} &= \mathbf{g}_{m, m_0}^{(L)} + \sum_{\alpha} \left(\sum_{\beta} e^{ik_{-, \beta}^{(L)}(m-m_L)a} |\Phi_{-, \beta}^{(L)}\rangle \mathcal{S}_{\beta, \alpha}^{(LL)} \right) \\ &\times e^{ik_{+, \alpha}^{(L)}(m_L-m_0)a} \langle \phi_{+, \alpha}^{(L)} | \mathbf{g}^{(L)}, \end{aligned} \quad (80c)$$

$$\begin{aligned} \mathbf{G}_{m \in R, m_0 \in R} &= \mathbf{g}_{m, m_0}^{(R)} + \sum_{\alpha} \left(\sum_{\beta} e^{ik_{+, \beta}^{(R)}(m-m_R)a} |\Phi_{+, \beta}^{(R)}\rangle \mathcal{S}_{\beta, \alpha}^{(RR)} \right) \\ &\times e^{ik_{-, \alpha}^{(R)}(m_R-m_0)a} \langle \phi_{-, \alpha}^{(R)} | \mathbf{g}^{(R)}. \end{aligned} \quad (80d)$$

The above expressions have clear physical meanings. For example, Eq. (80a) shows that a local excitation $|\Phi_{\text{loc}}\rangle_{m_0}$ in the left lead first evolves to a partial wave $\mathbf{g}^{(L)} | \Phi_{\text{loc}} \rangle_{m_0}$ and then propagates freely to the left surface of the scatterer as $\sum_{\alpha} e^{ik_{+, \alpha}^{(L)}(m_L-m_0)a} |\Phi_{+, \alpha}^{(L)}\rangle \langle \phi_{+, \alpha}^{(L)} | \mathbf{g}^{(L)} | \Phi_{\text{loc}} \rangle_{m_0}$. Finally each eigenmode $|\Phi_{+, \alpha}^{(L)}\rangle$ evolves to a transmission wave:

$$|\Phi_{+, \alpha}^{(L)}\rangle \rightarrow \sum_{\beta} e^{ik_{+, \beta}^{(R)}(m-m_R)a} |\Phi_{+, \beta}^{(R)}\rangle \mathcal{S}_{\beta, \alpha}^{(RL)}. \quad (81)$$

As another example, Eq. (80c) shows that $\mathbf{G}_{m \in L, m_0 \in L}$ is the sum of the free GF $\mathbf{g}_{m, m_0}^{(L)}$ and the reflection wave contribution, which emerges as follows: the local excitation $|\Phi_{\text{loc}}\rangle_{m_0}$ in the left lead first evolves to a partial wave $\mathbf{g}^{(L)} | \Phi_{\text{loc}} \rangle_{m_0}$ and then propagates freely to the left surface of the scatterer as $\sum_{\alpha} e^{ik_{+, \alpha}^{(L)}(m_L-m_0)a} |\Phi_{+, \alpha}^{(L)}\rangle \langle \phi_{+, \alpha}^{(L)} | \mathbf{g}^{(L)} | \Phi_{\text{loc}} \rangle_{m_0}$. Finally, each mode evolves to a reflection wave:

$$|\Phi_{+, \alpha}^{(L)}\rangle \rightarrow \sum_{\beta} e^{ik_{-, \beta}^{(L)}(m-m_L)a} |\Phi_{-, \beta}^{(L)}\rangle \mathcal{S}_{\beta, \alpha}^{(LL)}. \quad (82)$$

Equation (80) not only allows us to construct the GF from the generalized scattering matrix, but also reveals the contribution of each individual scattering channels to the GF.

C. On-shell spectral expansion

The Fisher-Lee relations^{14-17,25,41,55} and its inverse [Eq. (80)] provide a complete one-to-one correspondence between the lattice GF approach and the wave function mode matching approach⁵²⁻⁵⁵ to mesoscopic quantum transport (see Fig. 1). Next we show that we can construct the GF $\mathbf{G}(E)$ analytically in terms of a few scattering states on the energy shell E . Since the latter can be readily obtained from standard textbook technique and approximation methods (such as the WKB approximation), this provides a convenient way to obtain the GF.

Let us introduce $2M$ advanced eigenmodes $\{|\tilde{k}_{s, \alpha}\rangle, |\tilde{\Phi}_{s, \alpha}\rangle\}$ of each lead⁵⁵,

$$\begin{aligned} \tilde{k}_{s, \alpha} &\equiv k_{s, \alpha}, \quad |\tilde{\Phi}_{s, \alpha}\rangle \equiv |\Phi_{s, \alpha}\rangle \quad (\alpha \in \text{evanescent}), \\ \tilde{k}_{s, \alpha} &\equiv k_{-s, \alpha}, \quad |\tilde{\Phi}_{s, \alpha}\rangle \equiv |\Phi_{-s, \alpha}\rangle \quad (\alpha \in \text{traveling}), \end{aligned} \quad (83)$$

and their dual vectors:

$$\begin{bmatrix} \langle \tilde{\Phi}_{s, 1} | \\ \vdots \\ \langle \tilde{\Phi}_{s, M} | \end{bmatrix} \equiv \tilde{\mathbf{U}}_s^{-1}, \quad (84)$$

where $\tilde{\mathbf{U}}_s \equiv [|\tilde{\Phi}_{s, 1}\rangle, \dots, |\tilde{\Phi}_{s, M}\rangle]$. The advanced eigenvectors $\{|\tilde{\Phi}_{s, \alpha}\rangle\}$ and dual vectors $\{\langle \tilde{\Phi}_{s, \alpha} | \}$ obey exactly the same orthonormal and completeness relations as their retarded counterpart [Eqs. (24) and (25)]. Using⁵⁵

$$\mathbf{g}^{-1} = \sum_{\alpha} |\tilde{\Phi}_{-, \alpha}\rangle \frac{iv_{+, \alpha}}{a} \langle \phi_{+, \alpha} | = \sum_{\alpha} |\tilde{\Phi}_{+, \alpha}\rangle \frac{-iv_{-, \alpha}}{a} \langle \phi_{-, \alpha} |, \quad (85)$$

and the completeness relations Eq. (25), we obtain

$$\mathbf{g} = \sum_{\alpha} |\Phi_{+, \alpha}\rangle \frac{a}{iv_{+, \alpha}} \langle \tilde{\Phi}_{-, \alpha} | = \sum_{\alpha} |\Phi_{-, \alpha}\rangle \frac{a}{-iv_{-, \alpha}} \langle \tilde{\Phi}_{+, \alpha} | \quad (86)$$

$$\rightarrow \sum_{\alpha} |\Phi_{+, \alpha}\rangle \frac{a}{iv_{+, \alpha}} \langle \Phi_{+, \alpha} | = \sum_{\alpha} |\Phi_{-, \alpha}\rangle \frac{a}{-iv_{-, \alpha}} \langle \Phi_{-, \alpha} |, \quad (87)$$

where the second line holds when all the eigenmodes are traveling modes. Here a is the unit cell spacing of the lead and $v_{s, \alpha}$ is the generalized group velocity of the eigenmode (s, α) : for a traveling mode, it equals the group velocity [Eq. (21)]; for an evanescent mode, it is defined as

$$v_{+, \alpha} = v_{-, \alpha}^* = -ia \langle \Phi_{-, \alpha} | (e^{ik_{-, \alpha} a})^* \mathbf{t}^\dagger - e^{ik_{+, \alpha} a} \mathbf{t} | \Phi_{+, \alpha} \rangle. \quad (88)$$

Note that for an evanescent (traveling) mode, $v_{s, \alpha}$ depends (does not depend) on the choice of the phases of the eigenvectors $\{|\Phi_{s, \alpha}\rangle\}$.

To express the GF in terms of on-shell scattering states, we introduce the eigenmode wave functions

$$|\Phi_{+, \alpha}(m)\rangle \equiv \begin{cases} e^{ik_{+, \alpha}(m-m_0)a} |\Phi_{+, \alpha}\rangle & (m \geq m_0), \\ 0 & (m < m_0), \end{cases} \quad (89a)$$

$$|\Phi_{-, \alpha}(m)\rangle \equiv \begin{cases} 0 & (m \geq m_0), \\ e^{ik_{-, \alpha}(m-m_0)a} |\Phi_{-, \alpha}\rangle & (m < m_0), \end{cases} \quad (89b)$$

for the lead in which m_0 locates. Note that so-defined $|\Phi_{s, \alpha}(m)\rangle$ depends on m_0 , which is regarded as fixed and hence

omitted for brevity. If there were no scatterers, then \mathbf{G}_{m,m_0} would coincide with the free GF of this lead [Eq. (37)], which can be written as

$$\mathbf{g}_{m,m_0} = a \sum_{\alpha} \left(\frac{|\Phi_{+, \alpha}(m)\rangle\langle\tilde{\Phi}_{-, \alpha}|}{i\nu_{+, \alpha}} + \frac{|\Phi_{-, \alpha}(m)\rangle\langle\tilde{\Phi}_{+, \alpha}|}{-i\nu_{-, \alpha}} \right). \quad (90)$$

Due to the presence of scatterers, each eigenmode $|\Phi_{s, \alpha}(m)\rangle$ evolves to a corresponding scattering state $|\Psi_{s, \alpha}(m)\rangle$, so replacing $|\Phi_{s, \alpha}(m)\rangle$ in Eq. (90) with $|\Psi_{s, \alpha}(m)\rangle$ gives the GF:

$$\mathbf{G}_{m,m_0} = a \sum_{\alpha} \left(\frac{|\Psi_{+, \alpha}(m)\rangle\langle\tilde{\Phi}_{-, \alpha}|}{i\nu_{+, \alpha}} + \frac{|\Psi_{-, \alpha}(m)\rangle\langle\tilde{\Phi}_{+, \alpha}|}{-i\nu_{-, \alpha}} \right). \quad (91)$$

Since the total scattering state $|\Psi_{s, \alpha}(m)\rangle$ can be decomposed into the sum of the incident wave $|\Phi_{s, \alpha}(m)\rangle$ (which vanishes outside the lead in which m_0 locates) and the outgoing scattering wave $|\Psi_{s, \alpha}^{(\text{out})}(m)\rangle \equiv |\Psi_{s, \alpha}(m)\rangle - |\Phi_{s, \alpha}(m)\rangle$, Eq. (91) can also be written as

$$\mathbf{G}_{m,m_0} = \mathbf{g}_{m,m_0} + a \sum_{\alpha} \left(\frac{|\Psi_{+, \alpha}^{(\text{out})}(m)\rangle\langle\tilde{\Phi}_{-, \alpha}|}{i\nu_{+, \alpha}} + \frac{|\Psi_{-, \alpha}^{(\text{out})}(m)\rangle\langle\tilde{\Phi}_{+, \alpha}|}{-i\nu_{-, \alpha}} \right), \quad (92)$$

where \mathbf{g}_{m,m_0} is nonzero only for m in the same lead as m_0 .

Equation (91) or (92) shows that the GF $\mathbf{G}_{m,m_0}(E)$ is essentially certain scattering states $\{|\Psi_{s, \alpha}(m)\rangle\}$ (on the energy shell E) that obey outgoing boundary conditions; i.e., they emanate from the on-shell eigenmodes $|\Phi_{s, \alpha}(m)\rangle$ of the lead in which the local excitation occurs. Compared with the standard spectral expansion in classic textbook on quantum mechanics^{78–80} that expresses the GF in terms of all the eigenstates (both on-shell eigenstates and off-shell ones) of the system, Eq. (91) or (92) deepens our physical understanding about the GF and allows analytical reconstruction of the GF from a few scattering states of the system.

As an example, let us consider an infinite system containing a single scatterer (Fig. 4). For m_0 in the left lead, the left-going eigenmode $|\Phi_{-, \alpha}^{(L)}(m)\rangle$ is not scattered, so $|\Psi_{-, \alpha}^{(L, \text{out})}(m)\rangle = 0$, while the right-going eigenmode $|\Phi_{+, \alpha}^{(L)}(m)\rangle$ produces an outgoing scattering wave

$$|\Psi_{+, \alpha}^{(L, \text{out})}(m)\rangle|_{m \in L} = \left(\sum_{\beta} e^{ik_{-, \beta}^{(L)}(m-m_L)a} |\Phi_{-, \beta}^{(L)}\rangle \mathcal{S}_{\beta, \alpha}^{(LL)} \right) e^{ik_{+, \alpha}^{(L)}(m_L-m_0)a}, \quad (93)$$

$$|\Psi_{+, \alpha}^{(L, \text{out})}(m)\rangle|_{m \in R} = \left(\sum_{\beta} e^{ik_{+, \beta}^{(R)}(m-m_R)a} |\Phi_{+, \beta}^{(R)}\rangle \mathcal{S}_{\beta, \alpha}^{(RL)} \right) e^{ik_{+, \alpha}^{(L)}(m_L-m_0)a}. \quad (94)$$

Substituting into Eq. (92) gives Eqs. (80a) and (80c). Similarly, for m_0 in the right lead, the right-going eigenmode $|\Phi_{+, \alpha}^{(R)}(m)\rangle$ is not scattered, so $|\Psi_{+, \alpha}^{(R, \text{out})}(m)\rangle = 0$, while the left-

going eigenmode $|\Phi_{-, \alpha}^{(R)}(m)\rangle$ produces the outgoing wave

$$|\Psi_{-, \alpha}^{(R, \text{out})}(m)\rangle|_{m \in L} = \left(\sum_{\beta} e^{ik_{-, \beta}^{(L)}(m-m_L)a} |\Phi_{-, \beta}^{(L)}\rangle \mathcal{S}_{\beta, \alpha}^{(LR)} \right) e^{ik_{-, \alpha}^{(R)}(m_R-m_0)a}, \quad (95)$$

$$|\Psi_{-, \alpha}^{(R, \text{out})}(m)\rangle|_{m \in R} = \left(\sum_{\beta} e^{ik_{+, \beta}^{(R)}(m-m_R)a} |\Phi_{+, \beta}^{(R)}\rangle \mathcal{S}_{\beta, \alpha}^{(RR)} \right) e^{ik_{-, \alpha}^{(R)}(m_R-m_0)a}. \quad (96)$$

Substituting them into Eq. (92) gives Eqs. (80b) and (80d).

V. EXAMPLE AND APPLICATIONS

Here we first exemplify our general results in a 1D chain and then apply the formalism to graphene p - n junctions described by the tight-binding model.

A. Simple example: 1D chain

We consider a 1D chain characterized by one basis state in each unit cell, the unit cell Hamiltonian $\mathbf{h} = \varepsilon_0$, and the nearest-neighbor hopping $\mathbf{t} = -t < 0$. For a given wave vector k , Eq. (20) can be solved to yield the energy $E(k) = \varepsilon_0 - 2t \cos(ka)$, which is real in two cases: (1) $k \in \mathbb{R}$; (2) $k = ik$ or $k = \pi/a + ik$ with $\kappa \in \mathbb{R}$. The former gives the real energy band, while the latter gives the complex energy bands. For specificity we consider the energy $E \in [\varepsilon_0 - 2t, \varepsilon_0 + 2t]$, so there is one right-going traveling eigenmode $k_+ = k$ with group velocity $v = 2at \sin(ka) > 0$ and one left-going traveling eigenmode $k_- = -k$ with group velocity $-v$, where k is the positive solution to $E = E(k)$. The eigenvectors of the eigenmodes and dual vectors are $\Phi_{\pm} = \phi_{\pm} = 1$. The propagation matrices are $\mathbf{P}_{\pm} = e^{\pm ika}$, the conversion matrix of the lead is

$$\mathbf{g} = \frac{1}{2it \sin(ka)} = \frac{1}{iv/a}, \quad (97)$$

and the free GF of the lead is

$$\mathbf{g}_{m,m_0} = \frac{e^{ik|m-m_0|a}}{iv/a}. \quad (98)$$

When the on-site energy of the unit cell m_1 of the infinite 1D chain is replaced by $\varepsilon_0 + \delta$, the unit cell m_1 becomes a scatterer characterized the conversion matrix

$$\mathbb{G} = \frac{1}{iv/a - \delta}, \quad (99)$$

or equivalently the transmission amplitude $\mathcal{T} = \mathbb{G} \mathbf{g}^{-1} = (iv/a)/(iv/a - \delta)$ and reflection amplitude $\mathcal{R} = \mathbb{G} \mathbf{g}^{-1} - 1 = \delta/(iv/a - \delta)$. The GFs of the entire system are given by Eqs.

(45), (52), and (53) as

$$\mathbf{G}_{m \geq m_1, m_0 \leq m_1} = \frac{e^{ik(m-m_1)a} \mathcal{T} e^{ik(m_1-m_0)a}}{iv/a} = \mathcal{T} \frac{e^{ik(m-m_0)a}}{iv/a}, \quad (100)$$

$$\mathbf{G}_{m \leq m_1, m_0 \geq m_1} = \frac{e^{-ik(m-m_1)a} \mathcal{T} e^{-ik(m_1-m_0)a}}{iv/a} = \mathcal{T} \frac{e^{-ik(m-m_0)a}}{iv/a}, \quad (101)$$

$$\mathbf{G}_{m \leq m_1, m_0 \leq m_1} = \frac{e^{ika|m-m_0|}}{iv/a} + \frac{e^{-ik(m-m_1)a} \mathcal{R} e^{ik(m_1-m_0)a}}{iv/a}, \quad (102)$$

$$\mathbf{G}_{m \geq m_1, m_0 \geq m_1} = \frac{e^{ika|m-m_0|}}{iv/a} + \frac{e^{ik(m-m_1)a} \mathcal{R} e^{-ik(m_1-m_0)a}}{iv/a}. \quad (103)$$

The above results are also consistent with Eq. (80).

Finally, when the on-site energies of unit cell m_1 and unit cell m_2 ($> m_1$) are both replaced by $\varepsilon_0 + \delta$, then each unit cell becomes a scatterer characterized by the conversion matrix in Eq. (99). The conversion matrix $\mathbb{G}^{(C)}$ of the composite scatterer ($m_1 + m_2$) is given by Eq. (66). In particular, the surface elements are obtained from Eq. (67) as

$$\mathbb{G}_{m_2 m_1}^{(C)} = \mathbb{G}_{m_1 m_2}^{(C)} = e^{ik(m_2-m_1)a} \frac{\mathcal{T}(1 - e^{2ik(m_2-m_1)a} \mathcal{R}^2)^{-1} \mathcal{T}}{iv/a}, \quad (104)$$

$$\mathbb{G}_{m_1 m_1}^{(C)} = \mathbb{G}_{m_2 m_2}^{(C)} = \mathbb{G} + e^{2ik(m_2-m_1)a} \frac{\mathcal{T} \mathcal{R}(1 - e^{2ik(m_2-m_1)a} \mathcal{R}^2)^{-1} \mathcal{T}}{iv/a}, \quad (105)$$

from which we can obtain the generalized transmission amplitude across the composite scatterer as

$$\mathcal{T}^{(C)} \equiv \mathbb{G}_{m_2 m_1}^{(C)} \mathbf{g}^{-1} = \mathcal{T}(1 - e^{2ik(m_2-m_1)a} \mathcal{R}^2)^{-1} e^{ik(m_2-m_1)a} \mathcal{T}, \quad (106)$$

and the generalized reflection amplitude as

$$\mathcal{R}^{(C)} \equiv \mathbb{G}_{m_1 m_1}^{(C)} \mathbf{g}^{-1} - 1 = \mathcal{R} + e^{2ik(m_2-m_1)a} \mathcal{T} \mathcal{R}(1 - e^{2ik(m_2-m_1)a} \mathcal{R}^2)^{-1} \mathcal{T}. \quad (107)$$

B. Chiral tunneling and anomalous focusing in graphene p - n junction

Graphene is a single layer of carbon atoms in a honeycomb lattice that hosts massless Dirac fermions^{93–96}. One of the unique properties of electrons in graphene is chiral tunneling^{81–83}: an electron normally incident on a potential barrier will always be perfectly transmitted, independently of its kinetic energy and the height and width of the potential barrier (for oblique incidence, the transverse momentum serves as a gap-opening mass term, so the transmission is not perfect, and bound states may be created by a 1D potential well^{103,104}). In recent years, the chiral tunneling in graphene p - n junctions has attracted a lot of attention (see Ref. 97 for a review). Another interesting phenomenon for electrons in graphene p - n junctions is the anomalous focusing due to negative refraction, which was initially proposed by Veselago for electromagnetic waves^{98–101}: a spatially diverging pencil of rays is focused to a spatially converging one during the transition from a medium with positive refractive index across a sharp interface into a

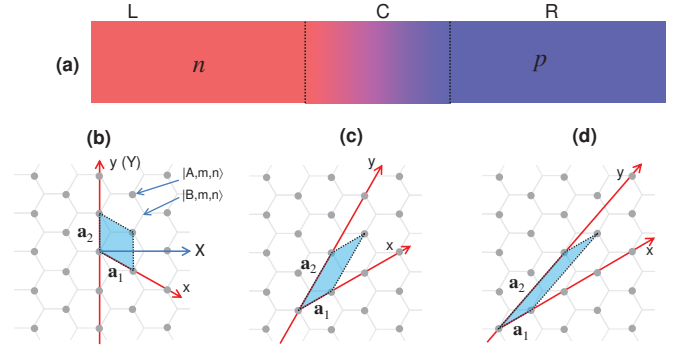


FIG. 7. (a) Sketch of graphene p - n junction with a smooth interface. Panels (b)-(d) show the choice of primitive vectors and x, y axes when the interface is along the zigzag direction (b), the armchair direction (c), or a more general direction (d). The filled dots mark the Bravais lattice (A sublattice) of graphene. The shaded regions mark the unit cells. In panel (b), the X and Y axes of the Cartesian coordinate system are also shown.

negative index medium. In 2007, Cheianov *et al.*⁸⁴ proposed that ballistic graphene p - n junctions can also exhibit negative refraction and hence focus the electron flow: in the electron-doped n (hole-doped p) region, the carrier group velocity is parallel (anti-parallel) to its momentum, in analogy to light propagation in a positive (negative) refractive index medium. Ever since then, there have been a lot of works on the negative refraction in graphene (see Refs. 85 and 86 for examples) and on the surface of topological insulators¹⁰². Recently, the Veselago lens effect in graphene was observed experimentally^{87,88}.

On the theoretical side, most of the previous studies focus on traveling states and are based on the low-energy continuous model, whose validity is limited to the vicinity of the Dirac points. A very recent work¹⁰⁵, based on the tight-binding model, calculates numerically the propagation of a wave packet in a large but finite graphene flake to approximate the Klein tunneling and caustics of electron waves. Here we apply our general GF formalism to study the chiral tunneling in an infinite graphene p - n junction and pay special attention to the *evanescent* eigenmode. Our approach provides a clear physical picture and allows us to calculate the GF over long distances, so we further perform both analytical analysis and numerical simulations of dual-probe STM measurements. Our results demonstrate the possibility of observing the spatially resolved interference pattern caused by the negative refraction in graphene p - n junctions and further reveal a few interesting features, such as the distance-independent conductance and its quadratic dependence on the carrier concentration, as opposed to the linear dependence in uniform graphene.

We consider a graphene p - n junction with an interface that can be either sharp or smooth, as shown in Fig. 7(a). In the tight-binding model, the interface could align along different crystalline directions. To provide the simplest description, a *tilted* coordinate system is usually necessary: we choose one primitive vector \mathbf{a}_2 (defined as the y axis) of the Bravais lattice of uniform graphene to be parallel to the interface, and choose the x axis of our tilted coordinate system to be parallel to the other primitive vector \mathbf{a}_1 ; i.e., the two nonorthogonal unit vec-

tors of the tilted coordinate are $\mathbf{e}_x \equiv \mathbf{a}_1/|\mathbf{a}_1|$ and $\mathbf{e}_y \equiv \mathbf{a}_2/|\mathbf{a}_2|$, as shown in Figs. 7(b)-7(d) for the interface along the zigzag direction, armchair direction, and a more general direction. This choice of the primitive vectors and the tilted coordinate system ensures that the lattice Hamiltonian is invariant upon translation by $|\mathbf{a}_2|$ along the y axis, so that the original 2D lattice model can be reduced to a 1D lattice model.

1. Reduction from 2D to 1D

For specificity, we assume that the interface is along the zigzag direction [Fig. 7(c)], where $|\mathbf{a}_1| = |\mathbf{a}_2| = \sqrt{3}a_{C-C} \equiv a$ and a_{C-C} is the C-C bond length. The vanishingly small spin-orbit coupling in graphene makes the GF spin-independent, so we neglect the electron spin. In the tight-binding model, each unit cell of graphene consists of $M = 2$ orbitals, i.e., $|A, m, n\rangle$ and $|B, m, n\rangle$ for the unit cell (m, n) , where m (n) is the index along the x (y) axis of the tilted coordinate and A, B labels the sublattice [see Fig. 7(c)]. The Hamiltonian $\hat{H} = \hat{H}_0 + \hat{V}$ is the sum of the uniform part

$$\hat{H}_0 = \sum_{m,n} t(|A, m+1, n\rangle + |A, m, n-1\rangle + |A, m, n\rangle) \langle B, m, n| + h.c. \quad (108)$$

and the p - n junction potential

$$\hat{V} = \sum_{m,n} V_m(|A, m, n\rangle \langle A, m, n| + |B, m, n\rangle \langle B, m, n|), \quad (109)$$

where $t \approx 3$ eV is the nearest-neighbor hopping constant⁹⁴. The junction potential V_m depends arbitrarily on m inside the interface ($m_L \leq m \leq m_R$), but equals V_L inside the left lead L ($m \leq m_L - 1$) and equals V_R inside the right lead R ($m \geq m_R + 1$).

Due to the invariance of \hat{H} upon translation by $|\mathbf{a}_2|$ along the y axis, the problem can be reduced from 2D to 1D by a Fourier transform

$$|k_y\rangle \equiv \frac{1}{\sqrt{N_y}} \sum_n e^{ik_y n a} |n\rangle, \quad |n\rangle \equiv \frac{1}{\sqrt{N_y}} \sum_{k_y} e^{-ik_y n a} |k_y\rangle, \quad (110)$$

where N_y is the number of unit cells along the y axis. In the new basis, the Hamiltonian $\hat{H} = \sum_{k_y} \hat{H}_{1D}(k_y) |k_y\rangle \langle k_y|$ is diagonal with respect to k_y , where $\hat{H}_{1D}(k_y)$ describes a 1D lattice with the unit cells labeled by m and each unit cell containing 2 basis states $|A\rangle, |B\rangle$. Let us use $\mathbf{R}_{mm} \equiv m\mathbf{a}_1 + n\mathbf{a}_2$ to denote the Bravais vector of the unit cell (m, n) and use $\mathbf{G}(\mathbf{R}_{m_2 n_2}, \mathbf{R}_{m_1 n_1}, E)$ to denote the retarded GF from the unit cell (m_1, n_1) to the unit cell (m_2, n_2) of the original 2D system, which is connected to the retarded GF $\mathbf{G}_{m_2, m_1}(E, k_y)$ of the 1D lattice from the unit cell m_1 to the unit cell m_2 via a Fourier transform

$$\mathbf{G}(\mathbf{R}_{m_2 n_2}, \mathbf{R}_{m_1 n_1}, E) = \frac{1}{N_y} \sum_{k_y} e^{ik_y(n_2 - n_1)a} \mathbf{G}_{m_2, m_1}(E, k_y), \quad (111)$$

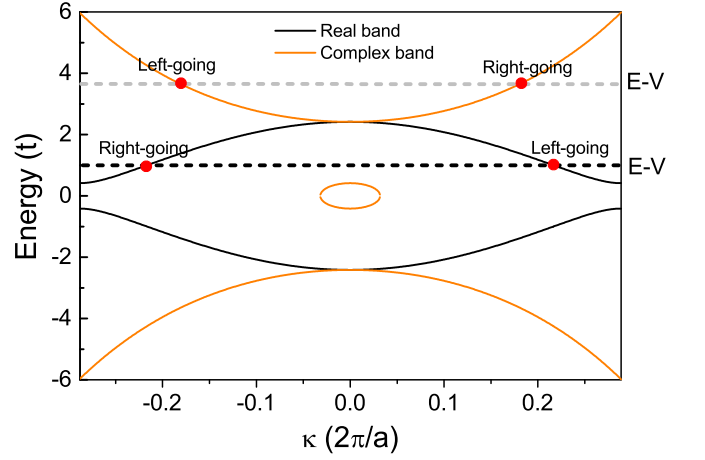


FIG. 8. Real energy band (black lines) and complex energy bands (orange lines) of pristine graphene at a fixed $k_y \approx 0.14 \times 2\pi/a$.

where $\mathbf{G}(\mathbf{R}_{m_2 n_2}, \mathbf{R}_{m_1 n_1}, E)$ and $\mathbf{G}_{m_2, m_1}(E, k_y)$ are both 2×2 matrices. Below we consider fixed E and k_y and apply our general results to calculate the GF \mathbf{G}_{m_2, m_1} of the 1D lattice, with E and k_y omitted for brevity.

2. Green's function of 1D lattice

In the 1D lattice, the hopping is uniform:

$$\mathbf{H}_{m, m+1}^{(1D)} = (\mathbf{H}_{m+1, m}^{(1D)})^\dagger = \mathbf{t} = \begin{bmatrix} 0 & 0 \\ t & 0 \end{bmatrix}. \quad (112)$$

The unit cell Hamiltonian $\mathbf{H}_{m, m}^{(1D)} = V_m + \mathbf{h}_0$ is the sum of the unit cell Hamiltonian of pristine graphene,

$$\mathbf{h}_0 = \begin{bmatrix} 0 & t(1 + e^{ik_y a}) \\ t(1 + e^{-ik_y a}) & 0 \end{bmatrix}. \quad (113)$$

and the p - n junction potential V_m . The entire infinite system consists of a single scatterer ($m_L \leq m \leq m_R$) connected to two semi-infinite leads L and R [cf. Fig. 7(a)], whose GFs can be constructed from the conversion and propagation matrices of the leads and the conversion matrix \mathbb{G} of the p - n interface (see Sec. III and Sec. IV).

The remaining issue is to calculate the eigenmodes of each lead, as characterized by the hopping \mathbf{t} and the unit cell Hamiltonian $\mathbf{h} = V + \mathbf{h}_0$, where $V = V_L$ (left lead) or V_R (right lead). Given a complex wave vector k , we can solve the eigenvalue problem Eq. (20) and obtain the energy bands of the lead as $V \pm E_0(k)$, where $E_0(k) \equiv t \sqrt{f(k, k_y) f(-k, -k_y)}$ and $f(k, k_y) \equiv 1 + e^{ik_y a} + e^{-ika}$. Here $E_0(k)$ is real in two cases: (1) $k + k_y/2 = \kappa$; (2) $k + k_y/2 = i\kappa$ or $\pi/a + i\kappa$, where $\kappa \in \mathbb{R}$. The former gives the real energy bands, while the latter gives the complex energy bands, as shown in Fig. 8. Conversely, given the energy E , we can solve Eq. (20) and obtain a right-going eigenmode k_+ , $|\Phi_+\rangle$ and a left-going eigenmode k_- , $|\Phi_-\rangle$, where k_\pm are the two solutions to $|E - V| = E_0(k)$ or equivalently the two intersection points of $E - V$ with the real

and complex energy bands of pristine graphene. When $E - V$ lies in the range of the real energy bands (dashed black line in Fig. 8), the wave vectors k_{\pm} are both real and both eigenmodes are traveling modes. When $E - V$ lies outside the range of the real energy bands (dashed gray line in Fig. 8), the wave vectors $k_{\pm} = k_{\pm}^*$ are complex and both eigenmodes are evanescent. A more convenient method to obtain the eigenmodes is to define $\lambda \equiv e^{ika}$ and rewrite Eq. (20) as

$$\lambda^2 bt + \lambda[|b|^2 + t^2 - (E - V)^2] + tb^* = 0 \quad (114)$$

with $b \equiv t(1 + e^{ik_y a})$, so that $\lambda_{\pm} = e^{ik_{\pm} a}$ are obtained as the two solutions to this quadratic equation for λ . In addition to the two *normal* eigenmodes, there are also a pair of *ideally evanescent* eigenmodes, including a right-going one $\lambda_{+,0} = e^{ik_{+,0} a} = 0$, $|\Phi_{+,0}\rangle = [1, 0]^T$ and a left-going one $\lambda_{-,0} = e^{ik_{-,0} a} = \infty$, $|\Phi_{-,0}\rangle = [0, 1]^T$ (see Appendix A). The ideally evanescent eigenmodes do not propagate, so they do not directly contribute to the GF, but their existence does influence the generalized transmission and reflection amplitudes of the normal eigenmodes.

Using the eigenmodes, we can calculate the conversion matrix \mathbb{G} of the p - n interface using Eq. (43) and then obtain the generalized transmission and reflection amplitudes $\mathcal{S}^{(LL)}$, $\mathcal{S}^{(RL)}$, $\mathcal{S}^{(LR)}$, and $\mathcal{S}^{(RR)}$ of the normal eigenmode from Eqs. (70)-(75). For $m \neq m_0$, the free GF of the left lead is essentially the sum of the left-going eigenmode $|\Phi_{-}^{(L)}(m)\rangle$ and the right-going eigenmode $|\Phi_{+}^{(L)}(m)\rangle$ [cf. Eq. (89) for their definitions]:

$$\mathbf{g}_{m,m_0}^{(L)} = \frac{a}{i\nu_{+}^{(L)}} |\Phi_{+}^{(L)}(m)\rangle \langle \tilde{\Phi}_{-}^{(L)}| + \frac{a}{-i\nu_{-}^{(L)}} |\Phi_{-}^{(L)}(m)\rangle \langle \tilde{\Phi}_{+}^{(L)}|. \quad (115)$$

Due to the p - n interface, the eigenmode $|\Phi_{+}^{(L)}(m)\rangle$ produces a reflection wave in the left lead and a transmission wave in the right lead:

$$|\Psi_{+}^{(L,\text{out})}(m)\rangle = \begin{cases} e^{ik_{-}^{(L)}(m-m_L)a} |\Phi_{-}^{(L)}\rangle \mathcal{S}^{(LL)} e^{ik_{+}^{(L)}(m_L-m_0)a} & (m \in L), \\ e^{ik_{+}^{(R)}(m-m_R)a} |\Phi_{+}^{(R)}\rangle \mathcal{S}^{(RL)} e^{ik_{+}^{(L)}(m_L-m_0)a} & (m \in R), \end{cases} \quad (116)$$

so the GF from the left lead to the right lead is essentially the transmission wave:

$$\mathbf{G}_{m \in R, m_0 \in L} = \frac{a}{i\nu_{+}^{(L)}} |\Psi_{+}^{(L,\text{out})}(m)\rangle \langle \tilde{\Phi}_{-}^{(L)}|, \quad (117)$$

while the GF inside the left lead is essentially the sum of the incident eigenmode $|\Phi_{\pm}^{(L)}(m)\rangle$ and the reflection wave:

$$\mathbf{G}_{m \in L, m_0 \in L} = \mathbf{g}_{m,m_0}^{(L)} + \frac{a}{i\nu_{+}^{(L)}} |\Psi_{+}^{(L,\text{out})}(m)\rangle \langle \tilde{\Phi}_{-}^{(L)}|. \quad (118)$$

3. Green's function of 2D graphene p - n junction: anomalous focusing

Compared with the standard RGF method, the advantages of our GF method lie in its physical transparency and numerical efficiency. Here we demonstrate the first point by using

our method to provide a clear physical picture of the anomalous focusing effect⁸⁴⁻⁸⁸ across the graphene p - n junction described by the tight-binding model. For this purpose, we first obtain the GF of the 2D graphene p - n junction from the 1D GFs by a Fourier transform [Eq. (111)]. In particular, the GF from the unit cell (m_1, n_1) in the n region to the unit cell (m_2, n_2) in the p region,

$$\mathbf{G}(\mathbf{R}_{m_2 n_2}, \mathbf{R}_{m_1 n_1}, E) = \int \frac{dk_y}{2\pi} |\Psi_{+}^{(L,\text{tran})}(\mathbf{R}_{m_2 n_2})\rangle \frac{a}{i\nu_{+}^{(L)}} \langle \tilde{\Phi}_{-}^{(L)}|, \quad (119)$$

is essentially the sum of all transmission wave functions

$$|\Psi_{+}^{(L,\text{tran})}(\mathbf{R}_{m_2 n_2})\rangle \equiv e^{i\mathbf{k}_{+}^{(R)} \cdot (\mathbf{R}_{m_2 n_2} - \mathbf{R}_{m_R 0})} |\Phi_{+}^{(R)}\rangle \mathcal{S}^{(RL)} e^{i\mathbf{k}_{+}^{(L)} \cdot (\mathbf{R}_{m_L 0} - \mathbf{R}_{m_1 n_1})}, \quad (120)$$

which emanates from the incident eigenmode $|\Phi_{+}^{(L)}(\mathbf{R}_{m,n})\rangle = e^{i\mathbf{k}_{+}^{(L)} \cdot (\mathbf{R}_{m,n} - \mathbf{R}_{m_1 n_1})} |\Phi_{+}^{(L)}\rangle$ through three steps: propagation to the left interface $\mathbf{R}_{m_L 0}$ of the junction with wave vector $\mathbf{k}_{+}^{(L)}$, transmission across the interface, and propagation from $\mathbf{R}_{m_R 0}$ to $\mathbf{R}_{m_2 n_2}$ with wave vector $\mathbf{k}_{+}^{(R)}$. Here $\mathbf{k}_{\pm}^{(L)}$ ($\mathbf{k}_{\pm}^{(R)}$) are the wave vectors of the *normal* eigenmodes in the left (right) region, i.e., in the tilted coordinate [Fig. 7(b)]: $\mathbf{k}_{\pm}^{(p)} \cdot \mathbf{e}_x = k_{\pm}^{(p)}$ and $\mathbf{k}_{\pm}^{(p)} \cdot \mathbf{e}_y \equiv k_y$ ($p = L, R$). The GF from $\mathbf{R}_{m_1 n_1}$ to $\mathbf{R}_{m_2 n_2}$ can be measured as the conductance between one STM probe coupled to $\mathbf{R}_{m_1 n_1}$ and another STM probe coupled to $\mathbf{R}_{m_2 n_2}$ through the Landauer-Büttiker formula¹ $(2e^2/h)T(E_F)$, where the transmission probability $T(E_F) \propto |\mathbf{G}(\mathbf{R}_{m_2 n_2}, \mathbf{R}_{m_1 n_1}, E_F)|^2$. Therefore, Eq. (119) directly connects the transmission wave function to the experimentally measurable conductance and hence provides a clear physical picture for observing the anomalous focusing in dual-probe STM measurements, and further reveals some interesting effects.

Let us consider a sharp, symmetric interface at $m_L = m_R = 0$ and $V_R = -V_L = V_0 > 0$. In this case, the transmission wave simplifies to

$$|\Psi_{+}^{(L,\text{tran})}(\mathbf{R}_{m_2 n_2})\rangle \equiv e^{i(\mathbf{k}_{+}^{(R)} \cdot \mathbf{R}_{m_2 n_2} - \mathbf{k}_{+}^{(L)} \cdot \mathbf{R}_{m_1 n_1})} |\Phi_{+}^{(R)}\rangle \mathcal{S}^{(RL)}. \quad (121)$$

Here $\mathbf{k}_{+}^{(R)}$, $\mathbf{k}_{+}^{(L)}$, $|\Phi_{+}^{(R)}\rangle$, and $\mathcal{S}^{(RL)}$ all depend on k_y weakly. The strongest dependence on k_y comes from the phase factor $e^{i(\mathbf{k}_{+}^{(R)} \cdot \mathbf{R}_{m_2 n_2} - \mathbf{k}_{+}^{(L)} \cdot \mathbf{R}_{m_1 n_1})}$, which usually oscillates rapidly as a function of k_y when $\mathbf{R}_{m_2 n_2}$ and $\mathbf{R}_{m_1 n_1}$ are far away. However, when the energy of the incident electron lies midway in between the Dirac point of the n region and the Dirac point of the p region (i.e., $E = E_F = 0$), in the *Cartesian* coordinate system spanned by the orthogonal unit vectors \mathbf{e}_x and \mathbf{e}_y [see Fig. 7(b)], the electron-hole symmetry of graphene dictates that the Fermi wave vector $\mathbf{k}_{+}^{(L)}$ of the right-going eigenmode in the n region and the Fermi wave vector $\mathbf{k}_{+}^{(R)}$ of the right-going eigenmode in the p region to have the same component along the p - n interface (i.e., $\mathbf{k}_{+}^{(L)} \cdot \mathbf{e}_y = \mathbf{k}_{+}^{(R)} \cdot \mathbf{e}_y$), but opposite components perpendicular to the p - n interface (i.e., $\mathbf{k}_{+}^{(L)} \cdot \mathbf{e}_x = -\mathbf{k}_{+}^{(R)} \cdot \mathbf{e}_x$). Therefore, when $\mathbf{R}_{m_1 n_1}$ and $\mathbf{R}_{m_2 n_2}$ are mirror symmetric about the p - n interface, the rapidly oscillating phase factor equals unity for all k_y . In this case, all the transmitted waves have nearly the same phase factor for all k_y , so they contribute constructively to the GF. This corresponds

to electrons flowing out of an electron source at $\mathbf{R}_{m_1 n_1}$ being focused to $\mathbf{R}_{m_2 n_2}$, i.e., the anomalous focusing⁸⁴. According to the Landauer-Büttiker formula, the constructive enhancement of the GF could be detected as an enhanced conductance in dual-probe STM measurements.

In addition to locally enhancing the GF, the constructive interference of all the transmission waves also gives rise to two interesting behaviors. First, the phase factor $e^{i(\mathbf{k}_+^{(R)} \cdot \mathbf{R}_{m_2 n_2} - \mathbf{k}_+^{(L)} \cdot \mathbf{R}_{m_1 n_1})}$ and hence the transmission wave and the GF remain invariant when $\mathbf{R}_{m_2 n_2}$ and $\mathbf{R}_{m_1 n_1}$ are moved equally but in opposite directions perpendicular to the p - n interface. This would give rise to distance-independent conductance. Second, since each transmission wave contributes constructively to the GF, we have $\mathbf{G}(\mathbf{R}_{m_2 n_2}, \mathbf{R}_{m_1 n_1}, E_F) \propto$ density of states on the Fermi surface $\propto V_0 \propto$ carrier concentration. Thus the locally enhanced conductance should increase quadratically with increasing doping level V_0 (or equivalently the carrier concentration), in contrast to the linear dependence in uniform graphene⁷¹. These points will be verified in our subsequent numerical simulations of the dual-probe STM measurements, which provide a useful tool, with high spatial resolution, to measure such local transport properties and detect possible zero-energy bound states of the Dirac fermions caused by suitable 2D potential well^{106,107}.

4. Comparison of the standard RGF and our GF approach

Compared with the standard RGF that treats the *entire* central region numerically [see Fig. 2(a) for an example], an important advantage of our approach is that it fully utilizes the translational invariance of all the periodic subregions [even if they lie inside the central region, such as the middle lead in Fig. 2(a)] to treat these subregions semianalytically, so that only the *truly disordered* subregions need numerical treatment. Therefore, our GF approach is more efficient if the central region contains periodic subregions; otherwise the two approaches are equally efficient. Here we calculate $G(\mathbf{R}_2, \mathbf{R}_1, E)$ across a sharp graphene p - n junction using these two methods to demonstrate their equivalence and highlight the numerical efficiency of our approach. In the calculation, \mathbf{R}_1 is fixed at a randomly chosen A-sublattice site in the n region; \mathbf{R}_2 is swept over all the B-sublattice sites along the x axis from the n region to the p region. The range of the sweep is from $0.25\mu\text{m}$ on the left of \mathbf{R}_1 to $0.75\mu\text{m}$ on the right of \mathbf{R}_1 . For the RGF method, the central region (infinite along the p - n interface) is the smallest region that encloses \mathbf{R}_1 , \mathbf{R}_2 , and the p - n interface. For our GF method, the scatterer region (infinite along the p - n interface) consists of one slice at the p - n interface. The numerical results from the two approaches always agree with each other up to the machine accuracy. The time cost, however, differs by two orders of magnitude: with 90 Intel cores, the time cost of the standard RGF approach varies from 260 s to 540 s, depending on the position of \mathbf{R}_1 relative to the p - n interface, while the time cost of our approach is always less than 3 s. Similar speedup is expected in long multilayer structures with sharp interfaces, such as quantum wells, superlattices, or sharp p - n junctions.

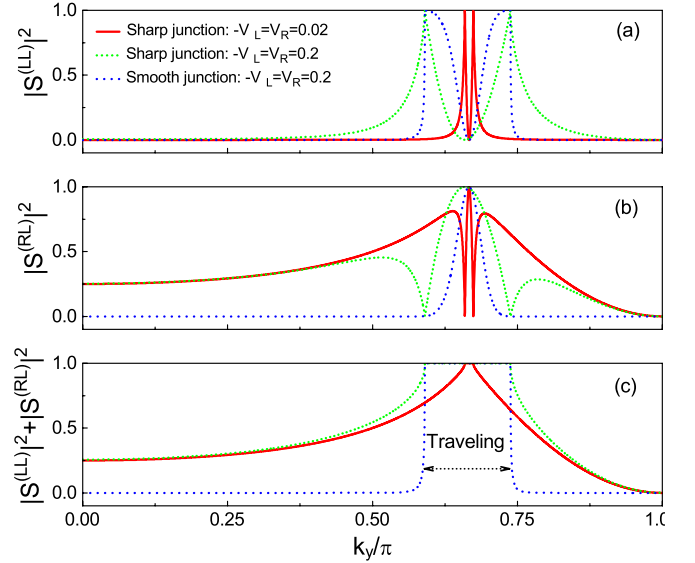


FIG. 9. Transmission and reflection of a right-going eigenmode of energy $E = 0$ incident from the n region of a sharp (solid lines and dotted lines) or 5nm wide smooth (dashed lines), symmetric (i.e., $-V_L = V_R = V_0$) graphene p - n junction. For $V_0 = 0.2$, the range of k_y in which the eigenmode is traveling is marked by the double arrow. For $V_0 = 0.02$, the range of k_y in which the eigenmode is traveling is much narrower.

5. Numerical examples

In the following (main text and all the figures), we always use the C-C bond length of graphene $a_{C-C} = 0.142$ nm as the unit of length and the nearest-neighbor hopping amplitude $t = 3$ eV as the unit of energy. For specificity, we focus on symmetric graphene p - n junctions with $V_R = -V_L = V_0 > 0$. Unless explicitly specified, we always take a typical doping level $V_0 = 0.2$ and set the energy $E = E_F = 0$, so we denote $G(\mathbf{R}_2, \mathbf{R}_1, E)$ by $G(\mathbf{R}_2, \mathbf{R}_1)$ for brevity.

As shown in Fig. 9, at $V_0 = 0.2$, the tunneling of a right-going traveling eigenmode reproduces the well-known results from the continuum model⁹⁷, such as the perfect transmission at normal incidence. For the evanescent eigenmode in a sharp p - n junction, however, $|S^{(RL)}|^2$ shows a peak, indicating enhanced tunneling of certain evanescent states. When the Fermi level is tuned closer to the Dirac point, i.e., for $V_0 = 0.02$, this enhanced tunneling becomes more pronounced and may be observed by dual-probe STM measurements.

Next we visualize the contributions from different scattering channels to the GF and their interference. For a smooth junction, the spatial map of the GF [Fig. 10(a)] shows imperfect focusing⁸⁴ due to negative refraction across a finite-width p - n junction. Let us consider $\mathbf{R}_{m_2 n_2}$ and $\mathbf{R}_{m_1 n_1}$ both in the n region and $\mathbf{R}_{m_2 n_2}$ on the right of $\mathbf{R}_{m_1 n_1}$ (i.e., $m_2 > m_1$); the GF

$$\begin{aligned} \mathbf{G}(\mathbf{R}_{m_2 n_2}, \mathbf{R}_{m_1 n_1}) = & \int \frac{dk_y}{2\pi} \frac{a}{iv_+^{(L)}} \\ & \times [|\Phi_+^{(L)}(\mathbf{R}_{m_2 n_2})\rangle + |\Psi_+^{(L, \text{refl})}(\mathbf{R}_{m_2 n_2})\rangle] \langle \tilde{\Phi}_-^{(L)} | \end{aligned} \quad (122)$$

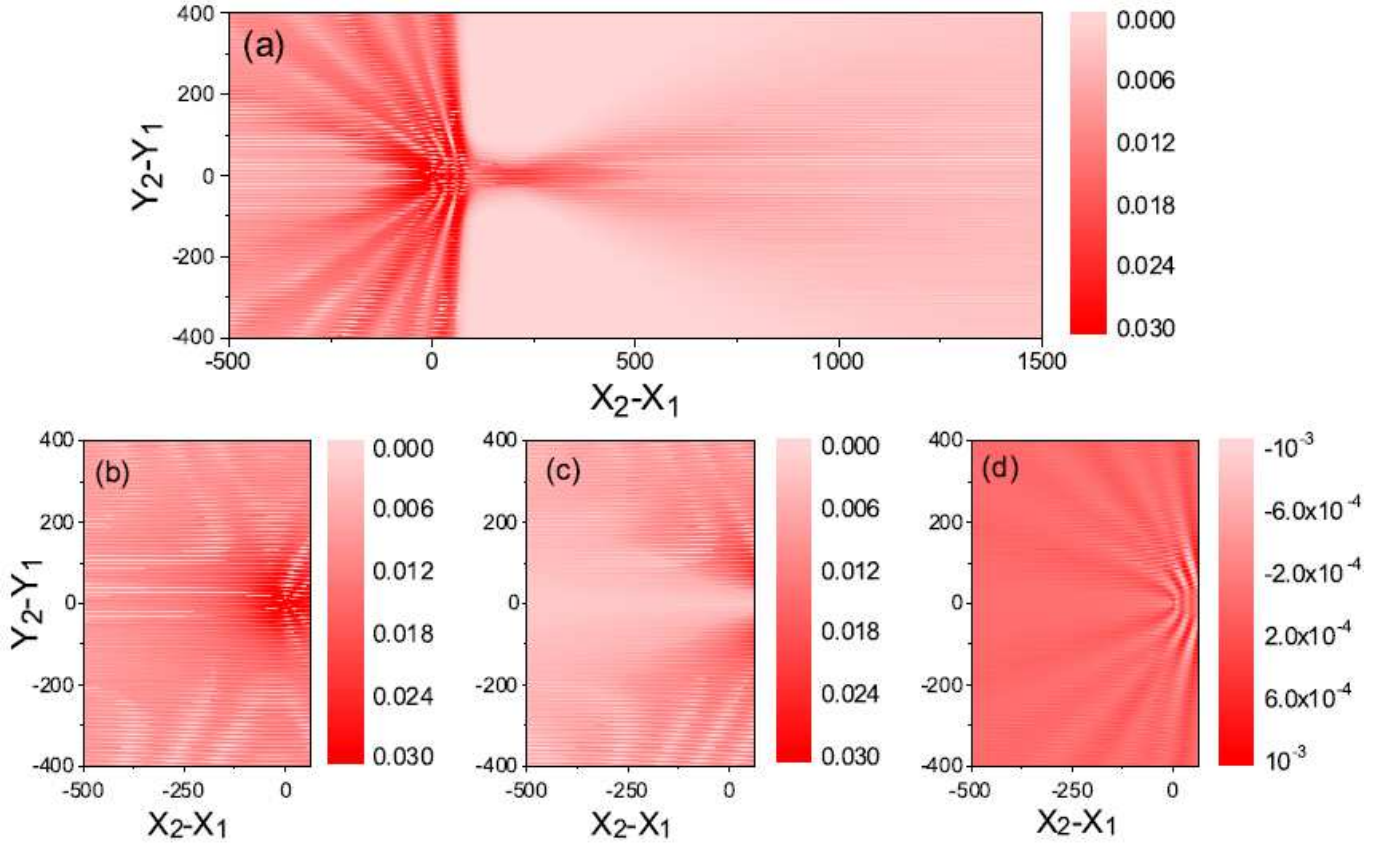


FIG. 10. Contributions of different scattering channels to the Green's function $G(\mathbf{R}_2, \mathbf{R}_1)$ for a smooth linear graphene p - n junction extending from 65 to 135 along the x axis (junction width ≈ 10 nm). Here $\mathbf{R}_1 \equiv (X_1, Y_1)$ is chosen to be fixed at $(0, 0)$ on the A sublattice, while $\mathbf{R}_2 \equiv (X_2, Y_2)$ is swept over all the B sublattice sites of the entire structure. (a) Contour plot of $|G(\mathbf{R}_2, \mathbf{R}_1)|$ vs $\mathbf{R}_2 - \mathbf{R}_1$. Panels (b) and (c) extract the contributions to $|G(\mathbf{R}_2, \mathbf{R}_1)|$ from the incident wave and reflection wave, respectively. Panel (d) shows the contribution to $|G(\mathbf{R}_2, \mathbf{R}_1)|^2$ due to the interference between the incident wave and the reflection wave.

is essential the sum of all right-going eigenmodes $|\Phi_+^{(L)}(\mathbf{R}_{m_2 n_2})\rangle$ and all reflection waves

$$|\Psi_+^{(L, \text{refl})}(\mathbf{R}_{m_2 n_2})\rangle \equiv e^{ik_-^{(L)} \cdot (\mathbf{R}_{m_2 n_2} - \mathbf{R}_{m_L 0})} |\Phi_-^{(L)}\rangle S^{(LL)} e^{ik_+^{(L)} \cdot (\mathbf{R}_{m_L 0} - \mathbf{R}_{m_1 n_1})}. \quad (123)$$

The incident wave contribution coincides with that of pristine graphene [Fig. 10(b)]. The reflection wave contribution [Fig. 10(c)] tends to vanish perpendicularly to the junction interface, indicative of the chiral tunneling. The interference between the incident and reflection waves [Fig. 10(d)] is responsible for the interference pattern in the total GF [Fig. 10(a)], which would be directly manifested in dual-probe STM measurements.

Finally we simulate dual-probe STM measurements^{71,108} over the graphene p - n junction at zero temperature. Following Refs. 71 and 108, we assume that each probe couples to a single carbon site for simplicity, so the Landauer-Büttiker formula¹ gives the interprobe conductance as $\sigma(\mathbf{R}_2, \mathbf{R}_1) = (2e^2/h) T_{12}(E_F) = \Gamma_1 \Gamma_2 |\bar{\mathbf{G}}(\mathbf{R}_2, \mathbf{R}_1, E_F)|^2$, where $\bar{\mathbf{G}}(\mathbf{R}_2, \mathbf{R}_1, E_F)$ is the GF incorporating the self-energy corrections from the STM probes and $\Gamma_{1,2}$ are coupling constants between the STM probes and the graphene. Usually, $\Gamma_{1,2}$ have a sensitive exponential dependence on the distance between the STM probe

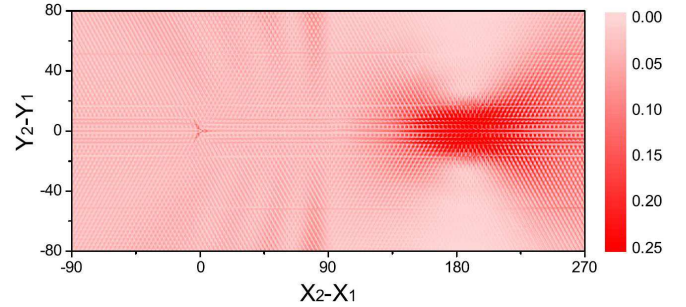


FIG. 11. Spatial map of the scaled conductance (i.e., the transmission coefficient T_{12} times $|\mathbf{R}_2 - \mathbf{R}_1|$) in a sharp graphene p - n junction as a function of $\mathbf{R}_2 - \mathbf{R}_1$. Here $\mathbf{R}_1 \equiv (X_1, Y_1)$ is fixed at the A sublattice of the unit cell $(0, 0)$ and $\mathbf{R}_2 \equiv (X_2, Y_2)$ is swept over all the lattice sites (including both A and B sublattices) of the entire structure.

and the graphene sample, but their specific values do not affect the shape of the signal. Therefore, following Ref. 108, we always rescale the maximum of T_{12} to unity. In Fig. 11, the real-space conductance map shows a pronounced focusing due to negative refraction⁸⁴. Other observable electron optics features include the high transparency of the junction near normal

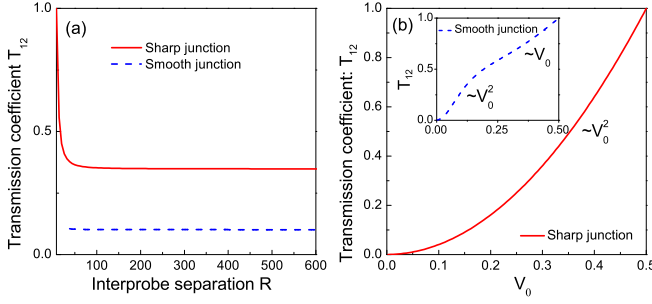


FIG. 12. Transmission coefficient T_{12} between two STM probes at mirror symmetric locations about the junction interface. (a) T_{12} vs. interprobe distance. (b) T_{12} vs. V_0 . Here, the width of the smooth junction is 5 nm for (a) and inset of (b), and the maximum of T_{12} is always rescaled to 1 in each panel.

incidence, i.e., chiral tunneling⁸², and the interference pattern between the incident and reflection waves. Recently, negative refraction in graphene p - n junctions was observed⁸⁷, but the measurement via macroscopic contacts only gives a spatially averaged result. Here our simulation shows that dual-probe STM measurements can further provide spatially resolved interference pattern; i.e., dual-probe STM could be an ideal experimental technique for studying local transport and quantum interference phenomenon.

Now we demonstrate numerically some interesting features of the dual-probe STM measurements in graphene p - n junctions [see the physical analysis following Eq. (121)]. First, when the two probes are mirror symmetric about the junction interface, T_{12} is nearly independent of the interprobe distance [Fig. 12(a)], in contrast to the $1/R$ decay in uniform graphene⁷¹. For a sharp junction, this behavior can be attributed to the anomalous focusing⁸⁴. However, the existence of the same behavior for a smooth junction interface indicates that it has a different physical origin, i.e., the cancellation of the propagation phases due to the matching of the electron Fermi surfaces in the n region and the hole Fermi surface in the p region¹⁰⁹. The distance-independent response could change qualitatively the spatial scaling of many physical quantities, such as the Friedel oscillation induced by an impurity and the carrier-mediated RKKY interaction between two localized magnetic moments. Second, the conductance across a sharp junction scales quadratically with the junction potential: $T_{12} \propto V_0^2$ [red line in Fig. 12(b)], which differs qualitatively from the linear scaling $T_{12} \propto V_0$ of uniform graphene⁷¹. When the junction becomes smooth [inset of Fig. 12(b)], the quadratic dependence still persists for small V_0 (electron's Fermi wavelength \gg junction width), but recovers the linear scaling behavior in uniform graphene for large V_0 . This can be attributed to the gradual destruction of the anomalous focusing when the junction width becomes larger than the Fermi wavelength⁸³.

VI. CONCLUSIONS

We have presented a numerically efficient and physically transparent formalism to calculate and understand the Green's function (GF) of a general layered structure. In contrast to the commonly used recursive GF method that directly calculates the GF through the Dyson equations, our approach converts the calculation of the GF to the generation and subsequent propagation of a scattering wave function emanating from a local excitation. This viewpoint provides analytical expressions of the GF in terms of a generalized scattering matrix. This identifies the contributions of individual scattering channels to the GF and hence allows this information to be extracted quantitatively from dual-probe STM experiments. We further derive an analytical rule to construct the GF of a general layered system, which could significantly reduce the computational time cost and enable quantum transport calculations for large samples. Application of this formalism to simulate the two-dimensional conductance map of a realistic graphene p - n junction demonstrates the possibility of observing the spatial interference caused by negative refraction and further reveals a few interesting features, such as the distance-independent conductance and its quadratic dependence on the carrier concentration, as opposed to the linear dependence in uniform graphene. In addition to conventional mesoscopic quantum transport, it would be interesting to apply our GF approach to the investigation other electron interference phenomena, such as the carrier-mediated RKKY interaction between local magnetic moments, the impurity-induced Friedel oscillation, and using dual-probe STM measurements to detect possible zero-energy bound states in graphene caused by suitable 2D potentials^{106,107}.

VII. ACKNOWLEDGMENTS

This work was supported by the MOST of China (Grants No. 2015CB921503 and No. 2014CB848700), the NSFC (Grants No. 11274036, No. 11322542, No. 11434010, and No.11504018), and the NSFC program for "Scientific Research Center" (Grant No. U1530401). We acknowledge the computational support from the Beijing Computational Science Research Center (CSRC).

Appendix A: Calculation of eigenmodes

When the determinant of the $M \times M$ hopping matrix \mathbf{t}^\dagger or \mathbf{t} is nonzero, the $2M$ eigenmodes can be obtained by letting $\lambda \equiv e^{ika}$ and rewriting Eq. (20) or equivalently $[-\mathbf{t}^\dagger + \lambda(z - \mathbf{h}) - \lambda^2 \mathbf{t}]|\Phi\rangle \equiv 0$ as a generalized $2M \times 2M$ eigenvalue problem

$$\begin{bmatrix} 0 & 1 \\ -\mathbf{t}^\dagger & z - \mathbf{h} \end{bmatrix} \begin{bmatrix} |\Phi\rangle \\ \lambda|\Phi\rangle \end{bmatrix} = \lambda \begin{bmatrix} 1 & 0 \\ 0 & \mathbf{t} \end{bmatrix} \begin{bmatrix} |\Phi\rangle \\ \lambda|\Phi\rangle \end{bmatrix}, \quad (\text{A1})$$

where $z = E + i0^+$. In the numerical calculation, we remove the infinitesimal imaginary part of z , i.e., $z = E$. Then Eq. (A1) is solved for $2M$ solutions $\{\lambda \equiv e^{ika}, |\Phi\rangle\}$, where

k could be either real (traveling modes) or complex (evanescent modes). Next the $2M$ eigenmodes are classified into M right-going ones and M left-going ones: the former consists of traveling modes (i.e., $\text{Im } k = 0$) with a positive group velocity and evanescent modes decaying exponentially to the right (i.e., $\text{Im } k > 0$), while the latter consists of traveling modes (i.e., $\text{Im } k = 0$) with a negative group velocity and evanescent modes decaying exponentially to the left (i.e., $\text{Im } k < 0$). There is an alternative but less accurate method to calculate and classify the eigenmodes. In this approach, the infinitesimal imaginary part of z is replaced by a finite but small positive number η , i.e., $z = E + i\eta$. Then Eq. (A1) is solved for the $2M$ solutions $\{\lambda \equiv e^{ika}, |\Phi\rangle\}$. Next, according to the sign of $\text{Im } k$, the $2M$ eigenmodes are classified into M right-going ones with $\text{Im } k > 0$ and M left-going ones with $\text{Im } k < 0$. In the limit $\eta \rightarrow 0^+$, the imaginary part $\text{Im } k$ may either vanish (i.e., traveling modes) or remain finite (i.e., evanescent modes). Obviously, the second approach is accurate only in the limit of $\eta \rightarrow 0^+$, so we always use the first approach in the main text.

When the determinant of the hopping matrix \mathbf{t} vanishes, solving Eq. (A1) would give some trivial evanescent eigenmodes. Suppose that M_0 out of the M eigenvalues of the hopping matrix \mathbf{t}^\dagger or \mathbf{t} are zero, and the corresponding eigenvectors are $\{|\Phi_{+, \alpha}\rangle\}$ (for \mathbf{t}^\dagger) and $\{|\Phi_{-, \alpha}\rangle\}$ (for \mathbf{t}), where $\alpha = 1, 2, \dots, M_0$. Then there would be $2M_0$ trivial evanescent eigenmodes, including M right-going ones $\lambda_{+, \alpha} = 0, |\Phi_{+, \alpha}\rangle$ and M_0 left-going ones $\lambda_{-, \alpha} = \infty, |\Phi_{-, \alpha}\rangle$. The former correspond to $|\Phi(m-1)\rangle = |\Phi_{+, \alpha}\rangle, |\Phi(m)\rangle = |\Phi(m+1)\rangle = 0$ in Eq. (19), while the latter correspond to $|\Phi(m+1)\rangle = |\Phi_{-, \alpha}\rangle, |\Phi(m)\rangle = |\Phi(m-1)\rangle = 0$ in Eq. (19). As a result, propagation matrices \mathbf{P}_- and \mathbf{P}_+^{-1} both diverge. However, this does not affect our formulas, which only contain \mathbf{P}_+ and \mathbf{P}_+^{-1} due to causality. The only problem is that for a trivial evanescent eigenmode (s, α) , the generalized group velocity $v_{s, \alpha}$ [Eq. (88)] is not well defined. For example, when $\lambda_{+, \alpha} = 0$ and hence $\lambda_{-, \alpha} = \infty$, the generalized velocity $v_{+, \alpha} = v_{-, \alpha}^* = -ia\lambda_{-, \alpha}^* \langle \Phi_{-, \alpha} | \mathbf{t}^\dagger | \Phi_{+, \alpha} \rangle$ involves the product of $\lambda_{-, \alpha}^* = \infty$ and $\mathbf{t}^\dagger | \Phi_{+, \alpha} \rangle = 0$. This singular problem can be avoided by adding sufficiently small numbers $\{\epsilon\}$ to \mathbf{t} and \mathbf{t}^\dagger to make their determinant nonzero and take the limit $\{\epsilon \rightarrow 0\}$ at the end of the calculation.

Taking the graphene junction as an example, the 1D left or right lead is characterized by the hopping matrix \mathbf{t} and unit cell Hamiltonian $V + \mathbf{h}_0$, where $V = V_L$ (left lead) or V_R (right lead) and \mathbf{h}_0, \mathbf{t} are given by Eqs. (112) and (113). Here the hopping matrix \mathbf{t} has one zero eigenvalue with eigenvector $|\Phi_{-, 0}\rangle = [0, 1]^T$, while \mathbf{t}^\dagger has one zero eigenvalue with eigenvector $|\Phi_{+, 0}\rangle = [1, 0]^T$. This gives rise to two trivial evanescent eigenmodes: the right-going one $\lambda_{+, 0} = 0, |\Phi_{+, 0}\rangle$ and the left-going one $\lambda_{-, 0} = \infty, |\Phi_{-, 0}\rangle$, for which the generalized group velocities $v_{+, 0} = v_{-, 0}^*$ are not well defined. To cure this problem, we add a small number ϵ to the off-diagonal of the hopping matrix, so that \mathbf{t} and \mathbf{t}^\dagger become

$$\mathbf{t}(\epsilon) = \begin{bmatrix} 0 & \epsilon \\ t & 0 \end{bmatrix}, \quad \mathbf{t}^\dagger(\epsilon) = \begin{bmatrix} 0 & t \\ \epsilon & 0 \end{bmatrix}. \quad (\text{A2})$$

Then using the first-order perturbation theory, we obtain

$$\lambda_{+, 0}(\epsilon) \approx \frac{-\epsilon}{t(e^{-iak_y} + 1)}, \quad |\Phi_{+, 0}(\epsilon)\rangle \approx \left[\frac{1}{t^2(e^{-iak_y} + 1)} \right], \quad (\text{A3})$$

$$\lambda_{-, 0}(\epsilon) \approx \frac{t(e^{iak_y} + 1)}{-\epsilon}, \quad |\Phi_{-, 0}(\epsilon)\rangle \approx \left[\frac{-\epsilon E}{t^2(e^{iak_y} + 1)} \right]. \quad (\text{A4})$$

Substituting into Eq. (88) and taking the limit $\epsilon \rightarrow 0$ gives

$$v_{+, 0} = v_{-, 0}^* = iat(e^{-iak_y} + 1). \quad (\text{A5})$$

Appendix B: Scattering of a partial wave

Let us consider a scatterer connected to a semi-infinite left lead (whose conversion matrix is \mathbf{g}) and prove that a right-going incident partial wave $|\Phi_{\text{in}}(m)\rangle$ is equivalent to a local excitation $|\Phi_{\text{loc}}\rangle_{m_0} \equiv \mathbf{g}^{-1}|\Phi_{\text{in}}(m_0)\rangle$ at an arbitrary site $m_0 \leq m_L$ (m_L is the left surface of the scatterer), in the sense that they produce the same scattering state at $m \geq m_0$. For this purpose, we use $|\Psi(m)\rangle$ for the conventional scattering state emanating from the incident wave $|\Phi_{\text{in}}(m)\rangle$ and $|\Phi(m)\rangle$ for the scattering state emanating from the local excitation $|\Phi_{\text{loc}}\rangle_{m_0}$ at m_0 . We notice that $|\Psi(m)\rangle$ and $|\Phi(m)\rangle$ obey the same Schrödinger equation for $m \geq m_0 + 1$, and the same uniform Schrödinger equation

$$-\mathbf{t}^\dagger |\Phi(m-1)\rangle + (z - \mathbf{h})|\Phi(m)\rangle - \mathbf{t}|\Phi(m+1)\rangle = 0 \quad (\text{B1})$$

for $m \leq m_0 - 1$. The difference lies at $m = m_0$:

$$-\mathbf{t}^\dagger |\Psi(m_0-1)\rangle + (z - \mathbf{H}_{m_0, m_0})|\Psi(m_0)\rangle - \mathbf{H}_{m_0, m_0+1}|\Psi(m_0+1)\rangle = 0, \quad (\text{B2})$$

$$-\mathbf{t}^\dagger |\Phi(m_0-1)\rangle + (z - \mathbf{H}_{m_0, m_0})|\Phi(m_0)\rangle - \mathbf{H}_{m_0, m_0+1}|\Phi(m_0+1)\rangle = |\Phi_{\text{loc}}\rangle_{m_0}, \quad (\text{B3})$$

and the boundary conditions: $|\Phi(m)\rangle$ should be finite at $m \rightarrow -\infty$, while the right-going part of $|\Psi(m)\rangle$ should equal the incident wave on the left of the scatterer, i.e., $|\Psi_+(m)\rangle = |\Phi_{\text{in}}(m)\rangle$ for $m \leq m_L$. The former gives $|\Phi(m_0-1)\rangle = \mathbf{P}_+^{-1}|\Phi(m_0)\rangle$, while the latter gives

$$|\Psi(m_0-1)\rangle = \mathbf{P}_+^{-1}|\Psi_+(m_0)\rangle + \mathbf{P}_-^{-1}|\Psi_-(m_0)\rangle \quad (\text{B4})$$

$$= \mathbf{P}_-^{-1}|\Psi(m_0)\rangle + (\mathbf{P}_+^{-1} - \mathbf{P}_-^{-1})|\Phi_{\text{in}}(m_0)\rangle. \quad (\text{B5})$$

Using these relations to eliminate $|\Phi(m_0-1)\rangle$ and $|\Psi(m_0-1)\rangle$ from the equations for $|\Phi(m)\rangle$ and $|\Psi(m)\rangle$ at $m = m_0$, and using Eq. (36), we see that they become identical and contains the same source $|\Phi_{\text{loc}}\rangle_{m_0}$. Therefore, $|\Phi(m)\rangle_{|m \geq m_0}$ and $|\Psi(m)\rangle_{|m \geq m_0}$ obeys exactly the same set of closed equations and natural boundary conditions (i.e., they should be finite for $m \rightarrow +\infty$); thus they are identical. Applying this equivalence principle to a scatterer connected to a semi-infinite left lead L and a semi-infinite right lead R gives Eqs. (46) and (49) of the main text.

Appendix C: Green's function of one scatterer: simple examples

Here we give a few simple examples for the GF of an infinite (or semi-infinite) system containing a single scatterer. As the first example, a semi-infinite lead (with unit cell Hamiltonian \mathbf{h} , hopping \mathbf{t} , and propagation matrices \mathbf{P}_\pm) consisting of the unit cells $m \leq 0$ can be regarded as a single-unit-cell scatterer at $m = 0$ connected to a semi-infinite left lead. Then the conversion matrix of this scatterer is

$$\mathbb{G}^{(L)} = (z - \mathbf{h} - \mathbf{t}^\dagger \mathbf{P}_-^{-1})^{-1} = (\mathbf{t} \mathbf{P}_-^{-1})^{-1}, \quad (\text{C1})$$

where we have used Eq. (36) in the second step. The GF of the entire system is

$$\mathbf{G}_{m,m_0}^{(L)} = \mathbf{g}_{m,m_0} + \mathbf{P}_-^m (\mathbb{G}^{(L)} \mathbf{g}^{-1} - \mathbf{I}) \mathbf{g}_{0,m_0}. \quad (\text{C2})$$

Taking $m = m_0 = 0$ gives $\mathbf{G}_{0,0}^{(L)} = \mathbb{G}^{(L)}$. By using Eqs. (C1) and (35), we have $\mathbb{G}^{(L)} \mathbf{g}^{-1} - \mathbf{I} = -\mathbf{P}_-^{-1} \mathbf{P}_+$ and hence recover the results by Sanvito *et al.*⁴¹: $\mathbf{G}_{m,m_0}^{(L)} = \mathbf{g}_{m,m_0} - \mathbf{P}_-^{m-1} \mathbf{P}_+^{1-m_0} \mathbf{g}$.

As the second example, a semi-infinite lead consisting of the unit cells $m \geq 0$ can be regarded as a single-unit-cell scatterer at $m = 0$ connected to a semi-infinite right lead. Then the conversion matrix of this scatterer is

$$\mathbb{G}^{(R)} = (z - \mathbf{h} - \mathbf{t} \mathbf{P}_+^{-1})^{-1} = (\mathbf{t}^\dagger \mathbf{P}_+^{-1})^{-1}. \quad (\text{C3})$$

The GF of the entire system is

$$\mathbf{G}_{m,m_0}^{(R)} = \mathbf{g}_{m,m_0} + \mathbf{P}_+^m (\mathbb{G}^{(R)} \mathbf{g}^{-1} - \mathbf{I}) \mathbf{g}_{0,m_0}, \quad (\text{C4})$$

where $\mathbb{G}^{(R)} \mathbf{g}^{-1} - \mathbf{I} = -\mathbf{P}_+ \mathbf{P}_-^{-1}$. Taking $m = m_0 = 0$ gives $\mathbf{G}_{0,0}^{(R)} = \mathbb{G}^{(R)}$.

The third example is an interface at $m = 0$ (with unit cell Hamiltonian $\mathbf{H}_{0,0}$) connected to two semi-infinite leads L and R . In this case the conversion matrix of the interface is

$$\mathbb{G}^{(I)} = (z - \mathbf{H}_{0,0} - \mathbf{t}_L^\dagger (\mathbf{P}_-^{(L)})^{-1} - \mathbf{t}_R \mathbf{P}_+^{(R)})^{-1}, \quad (\text{C5})$$

where \mathbf{t}_L (\mathbf{t}_R) is the nearest-neighbor hopping in the left (right) lead. The GFs of the entire system are given by

$$\mathbf{G}_{m \geq 0, m_0 \leq 0} = (\mathbf{P}_+^{(R)})^m \mathbb{G}^{(I)} (\mathbf{g}^{(L)})^{-1} \mathbf{g}_{0,m_0}^{(L)}, \quad (\text{C6})$$

$$\mathbf{G}_{m \leq 0, m_0 \geq 0} = (\mathbf{P}_-^{(L)})^m \mathbb{G}^{(I)} (\mathbf{g}^{(R)})^{-1} \mathbf{g}_{0,m_0}^{(R)}, \quad (\text{C7})$$

$$\mathbf{G}_{m \leq 0, m_0 \leq 0} = \mathbf{g}_{m,m_0}^{(L)} + (\mathbf{P}_-^{(L)})^m [\mathbb{G}^{(I)} (\mathbf{g}^{(L)})^{-1} - \mathbf{I}] \mathbf{g}_{0,m_0}^{(L)}, \quad (\text{C8})$$

$$\mathbf{G}_{m \geq 0, m_0 \geq 0} = \mathbf{g}_{m,m_0}^{(R)} + (\mathbf{P}_+^{(R)})^m [\mathbb{G}^{(I)} (\mathbf{g}^{(R)})^{-1} - \mathbf{I}] \mathbf{g}_{0,m_0}^{(R)}. \quad (\text{C9})$$

* wenyang@csrc.ac.cn

† kchang@semi.ac.cn

¹ S. Datta, *Electronic Transport in Mesoscopic Systems* (Cambridge University Press, Cambridge, England, 1995).

² D. Ferry and S. Goodnick, *Transport in Nanostructures* (Cambridge University Press, Cambridge, 1997).

³ Y. Blanter and M. Büttiker, *Phys. Rep.* **336**, 1 (2000).

⁴ A. Cresti, R. Farchioni, G. Grosso, and G. P. Parravicini, *Phys. Rev. B* **68**, 075306 (2003).

⁵ G. Metalidis and P. Bruno, *Phys. Rev. B* **72**, 235304 (2005).

⁶ R. Landauer, *IBM J. Res. Dev.* **1**, 223 (1957).

⁷ R. Landauer, *Philos. Mag.* **21**, 863 (1970).

⁸ M. Büttiker, Y. Imry, R. Landauer, and S. Pinhas, *Phys. Rev. B* **31**, 6207 (1985).

⁹ P. A. Khomyakov and G. Brocks, *Phys. Rev. B* **70**, 195402 (2004).

¹⁰ W. A. Harrison, *Electronic Structure and the Properties of Solids* (Dover, New York, 1989).

¹¹ J. Kudrnovský, V. Drchal, I. Turek, and P. Weinberger, *Phys. Rev. B* **50**, 16105 (1994).

¹² R. Zeller, P. H. Dederichs, B. Újfalussy, L. Szunyogh, and P. Weinberger, *Phys. Rev. B* **52**, 8807 (1995).

¹³ C. Caroli, R. Combescot, P. Nozieres, and D. Saint-James, *Journal of Physics C: Solid State Physics* **4**, 916 (1971).

¹⁴ D. S. Fisher and P. A. Lee, *Phys. Rev. B* **23**, 6851 (1981).

¹⁵ A. D. Stone and A. Szafer, *IBM J. Res. Develop.* **32**, 384 (1988).

¹⁶ H. U. Baranger and A. D. Stone, *Phys. Rev. B* **40**, 8169 (1989).

¹⁷ F. Sols, *Ann. Phys.* **214**, 386 (1992).

¹⁸ C. H. Lewenkopf and E. R. Mucciolo, *Journal of Computational Electronics* **12**, 203 (2013).

¹⁹ P. Drouvelis, P. Schmelcher, and P. Bastian, *Journal of Computa-*

tional Physics **215**, 741 (2006).

²⁰ D. J. Thouless and S. Kirkpatrick, *Journal of Physics C: Solid State Physics* **14**, 235 (1981).

²¹ P. A. Lee and D. S. Fisher, *Phys. Rev. Lett.* **47**, 882 (1981).

²² A. MacKinnon, *Zeitschrift für Physik B Condensed Matter* **59**, 385 (1985).

²³ H. U. Baranger, D. P. DiVincenzo, R. A. Jalabert, and A. D. Stone, *Phys. Rev. B* **44**, 10637 (1991).

²⁴ K. Kazymyrenko and X. Waintal, *Phys. Rev. B* **77**, 115119 (2008).

²⁵ M. Wimmer, Ph.D. thesis, University Regensburg (2009).

²⁶ X. Zheng, S.-H. Ke, and W. Yang, *J. Chem. Phys.* **132**, 114703 (2010).

²⁷ M. Settnes, S. R. Power, J. Lin, D. H. Petersen, and A.-P. Jauho, *Phys. Rev. B* **91**, 125408 (2015).

²⁸ A. MacKinnon and B. Kramer, *Phys. Rev. Lett.* **47**, 1546 (1981).

²⁹ C. M. Soukoulis, I. Webman, G. S. Grest, and E. N. Economou, *Phys. Rev. B* **26**, 1838 (1982).

³⁰ A. MacKinnon and B. Kramer, *Zeitschrift für Physik B Condensed Matter* **53**, 1 (1983).

³¹ L. Schweitzer, B. Kramer, and A. MacKinnon, *Journal of Physics C: Solid State Physics* **17**, 4111 (1984).

³² M. P. L. Sancho, J. M. L. Sancho, and J. Rubio, *Journal of Physics F: Metal Physics* **14**, 1205 (1984).

³³ M. P. L. Sancho, J. M. L. Sancho, and J. Rubio, *J. Phys. F: Met. Phys.* **15**, 851 (1985).

³⁴ H. Aoki and T. Ando, *Phys. Rev. Lett.* **54**, 831 (1985).

³⁵ T. Ando, *Phys. Rev. B* **40**, 5325 (1989).

³⁶ F. Sols, M. Macucci, U. Ravaioli, and K. Hess, *J. Appl. Phys.* **66**, 3892 (1989).

³⁷ T. Ando, *Phys. Rev. B* **42**, 5626 (1990).

- ³⁸ E. M. Godfrin, J. Phys. Condens. Matter **3**, 7843 (1991).
- ³⁹ M. B. Nardelli, Phys. Rev. B **60**, 7828 (1999).
- ⁴⁰ A. Umerski, Phys. Rev. B **55**, 5266 (1997).
- ⁴¹ S. Sanvito, C. J. Lambert, J. H. Jefferson, and A. M. Bratkovsky, Phys. Rev. B **59**, 11936 (1999).
- ⁴² F. Garcia-Moliner and V. Velasco, Surf. Sci. **299**, 332 (1994).
- ⁴³ J. Vasseur, A. Akjouj, L. Dobrzynski, B. Djafari-Rouhani, and E. E. Boudouti, Surf. Sci. Rep. **54**, 1 (2004).
- ⁴⁴ A. R. Rocha and S. Sanvito, Phys. Rev. B **70**, 094406 (2004).
- ⁴⁵ M. Brandbyge, J.-L. Mozos, P. Ordejón, J. Taylor, and K. Stokbro, Phys. Rev. B **65**, 165401 (2002).
- ⁴⁶ A. R. Rocha, V. M. García-Suárez, S. Bailey, C. Lambert, J. Ferrer, and S. Sanvito, Phys. Rev. B **73**, 085414 (2006).
- ⁴⁷ S. Birner, T. Zibold, T. Andlauer, T. Kubis, M. Sabathil, A. Trelakis, and P. Vogl, IEEE Trans. Electron Dev. **54**, 2137 (2007).
- ⁴⁸ T. Ozaki, K. Nishio, and H. Kino, Phys. Rev. B **81**, 035116 (2010).
- ⁴⁹ J. E. Fonseca, T. Kubis, M. Povolotskyi, B. Novakovic, A. Ajoy, G. Hegde, H. Ilatikhameneh, Z. Jiang, P. Sengupta, Y. Tan, et al., Journal of Computational Electronics **12**, 592 (2013).
- ⁵⁰ C. W. Groth, M. Wimmer, A. R. Akhmerov, and X. Waintal, New J. Phys. **16**, 063065 (2014).
- ⁵¹ S. Datta, Superlattices Microstruct. **28**, 253 (2000).
- ⁵² T. Ando, Phys. Rev. B **44**, 8017 (1991).
- ⁵³ K. Nikolić and A. MacKinnon, Phys. Rev. B **50**, 11008 (1994).
- ⁵⁴ K. Xia, M. Zwierzycki, M. Talanana, P. J. Kelly, and G. E. W. Bauer, Phys. Rev. B **73**, 064420 (2006).
- ⁵⁵ P. A. Khomyakov, G. Brocks, V. Karpan, M. Zwierzycki, and P. J. Kelly, Phys. Rev. B **72**, 035450 (2005).
- ⁵⁶ Note1, since $G(E)$ describes the scattering of both traveling and *evanescent* waves, while $S(E)$ only describes the scattering of traveling waves, $G(E)$ and $S(E)$ are equivalent only in the asymptotic regions of the lead, i.e., sufficiently far way from the scatterer such that the evanescent waves vanish completely.
- ⁵⁷ T. Nakayama, O. Kubo, Y. Shingaya, S. Higuchi, T. Hasegawa, C.-S. Jiang, T. Okuda, Y. Kuwahara, K. Takami, and M. Aono, Advanced Materials **24**, 1675 (2012).
- ⁵⁸ A.-P. Li, K. W. Clark, X.-G. Zhang, and A. P. Baddorf, Adv. Funct. Mater. **23**, 2509 (2013).
- ⁵⁹ O. Kubo, Y. Shingaya, M. Nakaya, M. Aono, and T. Nakayama, Appl. Phys. Lett. **88**, 254101 (2006).
- ⁶⁰ V. Cherepanov, E. Zubkov, H. Junker, S. Korte, M. Blab, P. Coenen, and B. Voigtlander, Rev. Sci. Instrum. **83**, 033707 (2012).
- ⁶¹ S. Qin, T.-H. Kim, Z. Wang, and A.-P. Li, Rev. Sci. Instrum. **83**, 063704 (2012).
- ⁶² H. Watanabe, C. Manabe, T. Shigematsu, and M. Shimizu, Appl. Phys. Lett. **78**, 2928 (2001).
- ⁶³ J. Baringhaus, M. Ruan, F. Edler, A. Tejeda, M. Sicot, Taleb-IbrahimiAmina, A.-P. Li, Z. Jiang, E. H. Conrad, C. Berger, et al., Nature **506**, 349 (2014).
- ⁶⁴ J. Baringhaus, M. Settnes, J. Aprojanz, S. R. Power, A.-P. Jauho, and C. Tegenkamp, Phys. Rev. Lett. **116**, 186602 (2016).
- ⁶⁵ P. W. Sutter, J.-I. Flege, and E. A. Sutter, Nat. Mater. **7**, 406 (2008).
- ⁶⁶ S.-H. Ji, J. B. Hannon, R. M. Tromp, V. Perebeinos, J. Tersoff, and F. M. Ross, Nat. Mater. **11**, 114 (2012).
- ⁶⁷ F. R. Eder, J. Kotakoski, K. Holzweber, C. Mangler, V. Skakalova, and J. C. Meyer, Nano Lett. **13**, 1934 (2013).
- ⁶⁸ K. W. Clark, X.-G. Zhang, I. V. Vlassiouk, G. He, R. M. Feenstra, and A.-P. Li, ACS Nano **7**, 7956 (2013).
- ⁶⁹ K. W. Clark, X.-G. Zhang, G. Gu, J. Park, G. He, R. M. Feenstra, and A.-P. Li, Phys. Rev. X **4**, 011021 (2014).
- ⁷⁰ T.-H. Kim, X.-G. Zhang, D. M. Nicholson, B. M. Evans, N. S. Kulkarni, B. Radhakrishnan, E. A. Kenik, and A.-P. Li, Nano Lett. **10**, 3096 (2010).
- ⁷¹ M. Settnes, S. R. Power, D. H. Petersen, and A.-P. Jauho, Phys. Rev. Lett. **112**, 096801 (2014).
- ⁷² S. Rotter, J.-Z. Tang, L. Wirtz, J. Trost, and J. Burgdörfer, Phys. Rev. B **62**, 1950 (2000).
- ⁷³ F. A. Maa, I. V. Zozulenko, and E. H. Hauge, Phys. Rev. B **50**, 17320 (1994).
- ⁷⁴ I. V. Zozulenko, F. A. Maa, and E. H. Hauge, Phys. Rev. B **53**, 7975 (1996).
- ⁷⁵ I. V. Zozulenko, F. A. Maa, and E. H. Hauge, Phys. Rev. B **53**, 7987 (1996).
- ⁷⁶ M.-H. Liu and K. Richter, Phys. Rev. B **86**, 115455 (2012).
- ⁷⁷ P. S. Krstić, X.-G. Zhang, and W. H. Butler, Phys. Rev. B **66**, 205319 (2002).
- ⁷⁸ J. J. Sakurai, *Modern Quantum Mechanics*, 2nd ed. (Addison-Wesley, San Francisco, 2001).
- ⁷⁹ D. J. Griffiths, *Introduction to Quantum Mechanics* (Prentice Hall, Upper Saddle River, New Jersey 07458, 1995).
- ⁸⁰ C. Cohen-Tannoudji, B. Diu, and F. Laloe, *Quantum Mechanics vol 2*, 2nd ed (Wiley-VCH, New York, 2005).
- ⁸¹ O. Klein, Z. Phys. **53**, 157 (1929).
- ⁸² M. I. Katsnelson, K. S. Novoselov, and A. K. Geim, Nat. Phys. **2**, 620 (2006).
- ⁸³ V. V. Cheianov and V. I. Fal'ko, Phys. Rev. B **74**, 041403 (2006).
- ⁸⁴ V. V. Cheianov, V. Fal'ko, and B. L. Altshuler, Science **315**, 1252 (2007).
- ⁸⁵ C.-H. Park, Y.-W. Son, L. Yang, M. L. Cohen, and S. G. Louie, Nano Lett. **8**, 2920 (2008).
- ⁸⁶ A. G. Moghaddam and M. Zareyan, Phys. Rev. Lett. **105**, 146803 (2010).
- ⁸⁷ G.-H. Lee, G.-H. Park, and H.-J. Lee, Nat. Phys. **11**, 925 (2015).
- ⁸⁸ S. Chen, Z. Han, M. M. Elahi, K. M. M. Habib, L. Wang, B. Wen, Y. Gao, T. Taniguchi, K. Watanabe, J. Hone, et al., Science **353**, 1522 (2016).
- ⁸⁹ Note2, lattice models with hopping between p th nearest neighbors can always be reduced to nearest-neighbor hopping by regarding p successive unit cells as one composite unit cell.
- ⁹⁰ J. Velev and W. Butler, J. Phys. Condens. Matter **16**, R637 (2004).
- ⁹¹ D. H. Lee and J. D. Joannopoulos, Phys. Rev. B **23**, 4997 (1981).
- ⁹² L. Molinari, J. Phys. A: Math. Gen. **30**, 983 (1997).
- ⁹³ C. W. J. Beenakker, Rev. Mod. Phys. **80**, 1337 (2008).
- ⁹⁴ A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, and A. K. Geim, Rev. Mod. Phys. **81**, 109 (2009).
- ⁹⁵ N. M. R. Peres, Rev. Mod. Phys. **82**, 2673 (2010).
- ⁹⁶ S. Das Sarma, S. Adam, E. H. Hwang, and E. Rossi, Rev. Mod. Phys. **83**, 407 (2011).
- ⁹⁷ P. E. Allain and J. Fuchs, The European Physical Journal B **83**, 301 (2011).
- ⁹⁸ V. G. Veselago, Soviet Physics Uspekhi **10**, 509 (1968).
- ⁹⁹ J. B. Pendry, Phys. Rev. Lett. **85**, 3966 (2000).
- ¹⁰⁰ X. Zhang and Z. Liu, Nat. Mater. **7**, 435 (2008).
- ¹⁰¹ J. B. Pendry, A. Aubry, D. R. Smith, and S. A. Maier, Science **337**, 549 (2012).
- ¹⁰² L. Zhao, P. Tang, B.-L. Gu, and W. Duan, Phys. Rev. Lett. **111**, 116601 (2013).
- ¹⁰³ J. M. Pereira, V. Mlinar, F. M. Peeters, and P. Vasilopoulos, Phys. Rev. B **74**, 045424 (2006).
- ¹⁰⁴ C. A. Downing and M. E. Portnoi, Phys. Rev. A **90**, 052116 (2014).
- ¹⁰⁵ R. Logemann, K. J. A. Reijnders, T. Tudorovskiy, M. I. Katsnelson, and S. Yuan, Phys. Rev. B **91**, 045420 (2015).
- ¹⁰⁶ J. H. Bardarson, M. Titov, and P. W. Brouwer, Phys. Rev. Lett. **102**, 226803 (2009).

- ¹⁰⁷ C. A. Downing, A. R. Pearce, R. J. Churchill, and M. E. Portnoi, Phys. Rev. B **92**, 165401 (2015).
- ¹⁰⁸ M. Settnes, S. R. Power, D. H. Petersen, and A.-P. Jauho, Phys. Rev. B **90**, 035440 (2014).
- ¹⁰⁹ S.-H. Zhang, J.-J. Zhu, W. Yang, H.-Q. Lin, and K. Chang, Phys. Rev. B **94**, 085408 (2016).